

Univerzita Karlova  
Filozofická fakulta

Fonetický ústav

*studijní program Filologie*  
*studijní obor Fonetika*

*Disertační práce*

Eliška Churaňová

**Temporální konfigurace českého taktu  
v souvislosti s jeho fonotaktickou strukturou**

**Temporal configurations of Czech stress-group  
in connection with its phonotactic structure**

2019

školitel: doc. PhDr. Jan Volín, Ph.D.

## **Poděkování**

Ráda bych tímto poděkovala svému školiteli doc. PhDr. Janu Volínovi, Ph.D. za provedení doktorským studiem a za rady, doporučení, inspiraci a směřování v rámci tvorby této práce.

Dále děkuji 40 dobrovolníkům, kteří absolvovali nikoli nenáročný percepční experiment, jenž tvořil jádro této studie.

Velký dík patří také mé rodině za pochopení a podporu v průběhu celého studia.

### **Prohlášení**

*Prohlašuji, že jsem disertační práci vypracovala samostatně, že jsem řádně citovala všechny použité prameny a literaturu a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.*

V Praze, dne

Podpis:

## Abstrakt

Tato disertační práce si klade za cíl prozkoumat vztahy mezi fonotaktickou stavbou českého mluvního taktu a řečovým rytmem. Pro účely této práce byly vybrány takty, jež odpovídaly třem nejčastějším dvouslabičným konsonanticko-vokálníkým vzorcům v češtině: CVCV, CCVCV a CVCCV. V percepčním experimentu, který zahrnoval srovnávání dvou plně proslovených mluvních taktů nebo taktu a nízkofrekvenčního obrazu jiného či stejného taktu, posluchači určovali, nakolik jsou si takty různých i stejných vzorců rytmicky podobné.

Výsledky ukázaly, že nejsilnější vliv na vnímání rytmické podobnosti má z fonotaktických faktorů pozice souhláskového shluku v mluvním taktu. O něco méně silnými faktory pro percepci řečového rytmu byly počet souhlásek v konsonantickém shluku a přítomnost dlouhého vokálu v obou porovnávaných taktech. Pro subtilnější rysy, jejichž význam pro percepci rytmické podobnosti prokázán nebyl (např. rozdíl v sonoritě znělých hlásek či znělosti obstruentů), byly navrženy možnosti dalšího zkoumání.

Klíčová slova: rytmus řeči, mluvní takt, fonotaktika, konsonanticko-vokálníká struktura, čeština

## **Abstract**

The main objective of this thesis is to explore the relationships between the phonotactic structure of the Czech stress-group and the rhythm of speech. Three most frequent consonantal-vocalic structures of Czech two-syllable stress-groups were selected for the purposes of this thesis: CVCV, CVCCV and CVCCV. In an auditory experiment, which contained the comparison of two stress-groups or the comparison of a stress-group and a low-frequency shadow of a stress-group, the respondents established how similar the rhythm of each couple of stress-groups sounded.

The results indicate that the position of a consonantal cluster within the stress-group is the strongest phonotactic factor in perception of the rhythmic similarity. The number of consonants within a consonantal cluster and the presence of a long vowel in both stress-groups were considered weaker factors for perceiving the rhythmic similarity by the respondents. Possibilities for a follow-up research were proposed for the factors that did not reach statistical significance, i.e., the difference in sonority or voicing of consonants.

**Keywords:** speech rhythm, stress-group, phonotactics, consonantal-vocalic structure, Czech

# Obsah

1 Úvod .....	8
1.1 Výzkumné otázky a cíle práce .....	8
1.2 Struktura práce.....	9
2 Teoretická východiska.....	10
2.1 Co je to rytmus .....	10
2.2 Zkoumání rytmického chování pomocí synchronizačních experimentů.....	11
2.3 Rytmus v řeči .....	12
2.4 Metody zkoumání rytmu řeči .....	13
2.5 O rytmu řeči v češtině .....	17
2.5.1 Faktory tvořící rytmické charakteristiky češtiny .....	18
2.5.1.1 Trvání hlásek v češtině .....	18
2.5.1.2 Slabika a výzkum p-center v češtině .....	22
2.5.1.3 Taktový přízvuk v češtině .....	24
2.5.1.4 Prozodická hierarchie souvislé češtiny .....	26
2.5.1.4.1 Mluvní takt.....	26
2.5.1.4.2 Promluvový úsek.....	29
2.5.1.4.3 Promluva .....	30
2.5.1.5 Větná prominence.....	31
2.5.1.6 Mluvní tempo a pauzy .....	32
2.5.1.6.1 Mluvní tempo.....	32
2.5.1.6.2 Pauzy .....	34
2.5.1.7 Mluvní styl .....	34
2.5.1.8 Fonotaktika češtiny .....	36
2.5.1.8.1 Stavba slabiky.....	36
2.5.1.8.2 Konsonanticko-vokalická stavba slabiky .....	39
2.5.1.8.3 Frekvence hlásek a jejich kombinací v češtině .....	41
2.5.1.8.4 Konsonanticko-vokalická stavba slov a mluvních taktů.....	47
3 Metoda.....	50
3.1 Materiál.....	50
3.1.1 Příprava materiálu .....	51
3.1.2 Fonotaktický popis materiálu.....	52
3.1.3 Materiál z perspektivy experimentu.....	54

3.2 Popis percepčního testu.....	57
3.2.1 Položky .....	57
3.2.2 Respondenti .....	59
3.2.3 Instrukce a průběh experimentu.....	59
3.2.4 Statistická analýza dat.....	60
4 Výsledky.....	61
4.1 Vlastnosti experimentálního plánu .....	61
4.1.1 Shoda nízkofrekvenčního obrazu s původním taktem.....	61
4.1.2 Přítomnost nízkofrekvenčního obrazu v položce.....	61
4.2 Fonotaktické vlastnosti a vnímání rytmu .....	62
4.2.1 Počet konsonantů v souhláskové skupině .....	62
4.2.2 Pozice souhláskového shluku v taktu .....	63
4.2.3 Vliv dlouhé samohlásky v taktu .....	64
4.2.4 Rozdíl v sonoritě znělých souhlásek .....	65
4.2.5 Kontrast znělosti .....	66
4.3 Suprasegmentální a biologické faktory .....	67
4.3.1 Vyhlazení intonačního průběhu.....	67
4.3.2 Rozdíl v pohlaví mluvčích v rámci jedné položky.....	67
4.4 Položková analýza .....	68
4.4.1 Obecné trendy v hodnocení položek.....	68
4.4.2 Kritéria výběru extrémně hodnocených položek .....	70
4.4.3 Vybrané položky.....	70
4.4.3.1 Položky s nízkým mediánem hodnocení .....	71
4.4.3.2 Položky s vysokým mediánem hodnocení.....	78
4.4.3.2.1 Položky obsahující rozdílné takty .....	78
4.4.3.2.2 Položky obsahující stejné takty .....	85
4.4.3.3 Položky s hodnocením obsahujícím více modů.....	88
5 Diskuse.....	92
6 Závěr.....	97
Literatura .....	99
Přílohy .....	110
Příloha 1 – položky testu.....	110
Příloha 2 – přepisy jednotlivých mluvních taktů v testu .....	115

# 1 Úvod

## 1.1 Výzkumné otázky a cíle práce

Tato disertační práce si klade za cíl pokračovat ve výzkumu nastíněném diplomovou prací téže autorky (Churaňová, 2012) a zasadit jeho závěry do širších souvislostí. Zatímco úkolem diplomové práce bylo popsat konsonanticko-vokální strukturu českého slova a mluvního taktu, cílem předkládaného výzkumu je popsat vztah mezi typickou fonotaktickou stavbou českého mluvního taktu a percepcí rytmu češtiny. Od zevrubného popisu, se kterými fonotaktickými vzorci v češtině lze počítat, se výzkum přesouvá k otázce, jak jsou tyto vzorce v jazyce vnímány jeho rodilými mluvčími.

Ačkoli je obecně přijímaným faktem, že uspořádání konsonantů a vokálů, z nichž jsou tvořeny vyšší rytmické jednotky, má vliv na percepci rytmu každého jazyka (např. Dauer, 1983), nebylo dosud v češtině stanoveno, jak konkrétní vlastnosti hlásek a jejich kombinací vnímání rytmičnosti ovlivňují. V rámci výzkumu, který je součástí této práce, bude tedy proveden komplexní průzkum vztahu mezi hláskovou či slabičnou strukturou slova a percepcí rytmu češtiny.

Mluvní takt je základní prozodickou jednotkou, na níž se realizuje rytmus řeči, a z tohoto důvodu byl zvolen jako výchozí jednotka i pro tento výzkum. V češtině existuje velké množství typů taktů, co se jejich konsonanticko-vokální struktury týče (např. CV, CCV, CVCV, CCVCV apod.), nicméně v projevech výrazně převažují takty dvouslabičné (Churaňová, 2012). V rámci této studie byly proto využity takty, které odpovídaly třem v češtině nejčastějším dvouslabičným konsonanticko-vokálním strukturám: CVCV, CCVCV a CVCCV.

Hypotéza počítala s řadou faktorů, jež by na vnímání rytmu mluvních taktů v češtině mohly mít vliv. Na základě výzkumu je možné konstatovat, které fonotaktické a hláskové činitele (např. počet konsonantů ve skupině souhlásek, pozice konsonantického shluku, znělost konsonantu apod.) se na percepci rytmu českých mluvních taktů podílejí více a které méně.

Faktory ovlivňující vnímání rytmu českých mluvních taktů byly stanoveny pomocí rozsáhlého percepčního experimentu. K odhalení vztahu mezi hláskami s jejich různými kombinacemi a rytmem přispěly jak obecná analýza kompletního testu, tak i podrobné prozkoumání jednotlivých položek v experimentu.



Pro percepční experiment byly využity nahrávky hlasatelů Českého rozhlasu, tzn. texty čtené profesionálními rodilými mluvčími češtiny. Tato forma mluveného projevu byla vybrána pro svou neutrálnost co do dialektu, idiolektu a expresivnosti, aby byla zajištěna srovnatelnost vlastností zkoumaných mluvních taktů a zároveň nebyla testována řeč laboratorní, s nejasnou či neúplnou komunikativní intencí.

## **1.2 Struktura práce**

V rámci teoretické části předkládané disertační práce je nejprve obecně pojednáno téma rytmu jakožto obecného jevu i rytmu v řeči včetně způsobů jeho zkoumání. Další kapitoly popisují současný stav poznání principů řečového rytmu v češtině s ohledem na souvislost s fonotaktickou stavbou jazyka.

V následujícím oddílu práce je popsán materiál pro hlavní experiment, jeho zpracování i průběh testu. Čtvrtá kapitola obsahuje kromě obecných závěrů plynoucích z vyhodnocení experimentu i analýzu jednotlivých položek percepčního testu, jejichž hodnocení respondenty bylo dle stanovených kritérií považováno za extrémní. V diskusi jsou pak nastíněny cesty, kterými by se mohl ubírat navazující výzkum.

## 2 Teoretická východiska

### 2.1 Co je to rytmus

Všechny formy rytmického chování, které se v přírodě objevuje v mnohých oblastech, vykazují jednu zásadní charakteristiku a tou je periodické opakování určitých vzorců. Tyto vzorce sestávají ze střídání kontrastů, které je recipientem vnímáno jako pravidelné. Jelikož je pravidelnost v produkci i percepci jednodušší než nepravidelnost, lze považovat rytmické procesy obecně za méně náročné a stabilnější, a tudíž i více žádoucí než arytmičné. Koncept stability byl v souvislosti s rytmem několikrát zkoumán na záměrné synchronizaci pohybů končetin s metronomovými pulzy ve zrychlujícím se tempu: experimenty shodně prokázaly, že od určité frekvence kliků se synchronizace přesunula z méně stabilního vzorce do stabilnějšího, například synchronizace v protifázi se změnila na koordinaci ve shodné fázi (Kelso, 1995; Repp, 2005).

Důležitost rytmických událostí je možné nalézt i na úrovni mozkového zpracování signálu: Model *neurální rezonance* (Grossberg, 2003), který se zabývá způsobem zpracování informace a jehož základem je kognitivní kódování v mozku, ukázal, že neuronová síť je schopna do jisté míry rozpoznat předcházející a následující vzorce. Série synchronizovaných aktivací neuronů se objeví, když se reprezentace vstupního signálu setkají s očekávanými neurálními reprezentacemi – neurony se tedy aktivují, když je očekávaný objekt ve shodě se vstupním signálem. V případě, že je rytmus nepředvídatelný, musejí neurony svoji činnost opakovat, dokud k synchronizované aktivaci (rezonanci) nedojde. Synchronizace akcí, typická pro rytmus, může probíhat jak v rámci daného organismu, tak i mezi různými organismy: Kelso (1995) tuto skutečnost považuje za projev dynamických procesů, jež operují na různých úrovních uvnitř jednoho systému i v součinnosti se systémy jinými.

Rytmické chování můžeme pozorovat na řadě živých organismů. Projevy rytmu spatřujeme i v samotné fyziologii množství živočichů – můžeme zmínit například pravidelnost dýchání, tlukot srdce, mávání křídlů při letu ptáků, pravidelné pohyby ploutví ryb a mořských savců. Nezastupitelnou roli má rytmus také v lidské kultuře: společenské aktivity, jako jsou hudba nebo tanec, jsou na vnímání, produkování či napodobování rytmických vzorců založeny.

## 2.2 Zkoumání rytmického chování pomocí synchronizačních experimentů

Jedním z nejpoužívanějších paradigmat, pomocí něhož je rytmus zkoumán, je tzv. *senzomotorická synchronizace*, kdy je akce koordinována s externím periodickým prediktabilním referentem, například metronomovým pulzem. Přehled dosavadních poznatků získaných od začátku 20. století pomocí tohoto paradigmatu, které zpravidla obsahuje vytūkávání prstem nebo pohyb, podává Bruno Repp (2005). Jak autor poukazuje, existují rozdílná provedení takového výzkumu – odlišnosti mohou být v typu akce (zkoumání vytūkávání prstem, nebo jiného pohybu), ve druhu koordinace (senzomotorická synchronizace vs. pokračovací/předcházející vytūkávání), ve způsobu synchronizace (vytūkávání ve fázi s podnětem nebo v protifázi; v poměru 1:1 či 1:n, n:1 apod.), v typu stimulů (vizuální, zvukové, manipulace s hmatovou odezvou). Při zkoumání synchronizace v hudbě se může také jednat o hledání dob, s kterými participant dokáže při poslechu hudby vytūkávat, nebo o synchronizaci několika hudebníků či hudebníka s nahrávkou. V neposlední řadě lze rytmus zkoumat pomocí studia neurálních korelátů senzomotorické synchronizace měřením mozkové aktivity pomocí funkční magnetické rezonance (fMRI), pozitronové emisní tomografie (PET), magnetoencefalografie (MEG), elektroencefalografie (EEG) nebo evokovaných potenciálů (ERP).

Repp (2005) ve svém souhrnu ukazuje několik aspektů typických pro produkci senzomotorické synchronizace. V první řadě se jedná o anticipační chování – po adaptaci na rytmus pulzů se participantova ťuknutí realizují před referenčními zvuky. Dále je potřeba vzít v úvahu, že jiné vytūkávání než ve fázi nebo v protifázi je pro testovanou osobu velmi nestabilní, přičemž i protifázové vytūkávání je náročnější – tento fakt potvrzují také zjištění Kelsa (1995). Běžný respondent, který není aktivním hudebníkem, je navíc schopen vyprodukovat maximálně 5–7 ťuknutí za sekundu (tzn. jeden úhoz za 150–200 ms), a pokud je frekvence pulzů vyšší, synchronizace se obvykle reorganizuje do způsobu vytūkávání 1:n (úhoz na každý n-tý pulz).

Jedním z nejpozoruhodnějších aspektů senzomotorické synchronizace je fakt, že není možné udržet koordinaci bez korekce chyb. Jak vyplývá z perturbačních studií (např. Repp & Keller, 2004), korekce rytmických chyb se pravděpodobně dá rozdělit do dvou odlišných procesů: korekce fáze, která je automatická a velmi rychlá, a korekce periody, která je kognitivně náročnější.

Zkoumání neurálních korelátů senzomotorické synchronizace (SMS) pomocí fMRI ukázalo, že vytukávání bez metronomu je náročnější než s ním: aktivují se přídatné neurální okruhy a mozková centra vykazují vyšší aktivitu. Rozdíly v mozkové aktivitě se rovněž projevily při fázovém a protifázovém vytukávání (Rao *et al.*, 1997; Jantzen *et al.*, 2002).

Senzomotorická synchronizace je paradigmatickým, které poskytuje vhodné podmínky pro zkoumání rytmu lidské řeči. Je možné experimentovat např. s rytmickým napodobením řečené výpovědi pomocí kontaktního vytukávání přízvukné slabiky či všech slabik nebo s vyslovováním daných promluv s metronomem. Pozornost lze věnovat i zkoumání asociace rytmických schopností v mluveném projevu s hudebními dovednostmi účastníka.

## 2.3 Rytmus v řeči

Rytmické aspekty nalezneme také v přirozené řeči. Střídající se opozice, které tvoří základ vnímání rytmičnosti, jsou tvořeny převážně jádry a hranami slabik nebo přízvuknými a nepřízvuknými slabikami. Přízvukné slabiky v projevu obvykle charakterizují odlišnosti od ostatních kvalitativně, kvantitativně či temporálně; přesto však řečový rytmus nelze popsat pouze pomocí jednoznačných objektivních kritérií.

Ve 40. letech 20. století byl představen rytmus řeči jakožto jeden z možných způsobů klasifikace rozdílných jazyků či jazykových rodin. Ačkoli Pike (1945) ve své práci nepovažoval principy taktové a slabičné izochronie za jednoznačná kritéria, která by měla jazyky rozčlenit do úhledně oddělených tříd – naopak sám upozorňoval na fakt, že jazyky obvykle uplatňují oba principy a liší se pouze v míře, kterou preferují izochronii zaměřenou na takty, nebo na slova – v následujících desetiletích ke zkoumání rytmu bylo přistupováno právě takto, jako by rytmus byl zcela objektivizovatelným faktorem, jehož se lze domoci exaktními postupy zkoumání řečových projevů. Z dnešní perspektivy je možné konstatovat, že tento způsob výzkumu se od pochopení podstaty rytmu řeči spíše vzdálil. Když Lehistová (1977, 1979) poukázala na to, že hodnocení rytmičnosti závisí na percepci, a nikoli jen na akustických charakteristikách stimulů, bylo nutné do koncepce zkoumání rytmu řeči zakomponovat posluchače. Tento trend výzkumu se v posledním desetiletí rozšířil nejen o vnímání rytmu, ale i o jeho komunikační funkci (Kohler, 2009). V současnosti je řečový rytmus interpretován jako hierarchicky uspořádaný jev, který

reflektuje principy organizace a periodické oscilace podobné té, která zakládá pohyb typický pro některé automatické úkony lidského těla nebo chování (Cummins, 2009), a různé oscilační vzorce odrážejí rytmus specifický pro rozdílné jazyky (Port, 2003).

Jak už bylo nastíněno, v rámci rytmických aspektů řeči je základním organizačním principem izochronie – časové intervaly mezi kontrastními řečovými jevy (např. přízvuky, jádra slabik, mory) jsou vnímány jako zhruba stejně dlouhé. Není zde ale řeč o čase objektivním, fyzikálním, nýbrž o čase subjektivně vnímaném, psychickém. Jedná se de facto o percepční iluzi, na niž upozornila Lehistová (1977, 1979) – řečové stimuly jsou totiž vnímány pravidelněji, než ve skutečnosti jsou. Toto tvrzení experimentálně potvrdili Donovan a Darwin (1979), kteří také zjistili, že mezi globální ukazatele rytmu patří kromě tempa i např. intonace, jejíž akustický korelát, základní frekvence, ovšem taktéž nemá přímý vztah k její percepci.

Rytmus řeči má poměrně velký vliv na percepci posluchače, a to nejen v případě významů slov a srozumitelnosti projevu, ale i co se týče toho, jak mluvčí na posluchače působí. Výše uvedený model neurální rezonance (Grossberg, 2003) poukázal i na důležitost rytmičnosti mluveného projevu – nepředvídatelný rytmus projevu má za následek vyšší kognitivní zátěž pro posluchače, což může ovlivnit pochopení sdělované informace i ochotu naslouchat dále.

Řečový rytmus není pro všechny jazyky stejný. Co je ale na první pohled méně zřetelné, je fakt, že existuje poměrně velká variabilita i v rámci jednoho jazyka – rytmus řeči se může měnit s ohledem na mluvní styl, tempo, komunikační situaci, recipienta i idiolektické rysy typické pro každého mluvčího. I přesto je ale podle všeho rytmičnost pro řeč přirozená: člověku obvykle nečiní velké potíže synchronizovat svůj projev se stejným projevem jiného mluvčího (Cummins, 2009).

## 2.4 Metody zkoumání rytmu řeči

Řečový rytmus jako samostatná charakteristika suprasegmentální roviny daného jazyka byl systematicky zkoumán od poloviny 20. století. Rozdíly ovšem nalezneme v optice, kterou tento fenomén jednotliví výzkumníci nahlíží. Podle Kohlera (2009) byla pro rané studie z 50. let typická perspektiva z hlediska **symbolické reprezentace** rytmu, která počítala se dvěma jeho aspekty: s akcentuací jakožto jazykovým rysem, který činí významné prvky prominentnějšími, a s rytmickým časováním coby charakteristikou,

která charakterizuje celé jazyky (např. s lexikálním přízvukem a větným akcentem). Také byla stanovena univerzální hierarchie akustických vlastností pro signalizování přízvuku:  $f_0$ , trvání, spektrum a energie (Fry, 1955, 1958).

Koncept slovního přízvuku ovlivnil uvažování o rytmu na dlouhou dobu – v anglofonním prostředí se již od 18. století soudilo, že přízvukné slabiky v angličtině se objevují v pravidelných intervalech, kteréžto přesvědčení vyústilo v rozdělení jazyků na slabičné a taktově (popř. morově) izochronní (Pike, 1945; Abercrombie, 1967). Obecný předpoklad zněl, že zatímco v tzv. *slabičně izochronních* (syllable-timed) jazycích trvají stejnou dobu intervaly mezi jednotlivými jádry slabik, v jazycích *taktově izochronních* (stress-timed) měly být co do trvání stejné intervaly mezi jednotlivými přízvuky. Abercrombie (1964) se snažil rytmickou jazykovou typologii definovat fyziologicky: každá slabika ve slabičně izochronních jazycích a každá přízvukná slabika v jazycích taktově izochronních měla být doprovázena hrudním pulzem, avšak pozdější výzkumy jeho předpoklady neprokázaly (Ladefoged, 1967; Ohala, 1975). Podle Dauerové (1983) měly taktově izochronní jazyky na rozdíl od slabičně izochronních také inklinovat k redukci vokálů v nepřízvukných pozicích a ke tvoření slabik se složitými konsonantickými skupinami.

V 70. a 80. letech 20. století byla izochronie reinterpretována jako **percepční** jev nebo tendence – sklon posluchačů percepčně podceňovat dlouhé intervaly a přeceňovat krátké. Při zkoumání percepčního základu izochronie se stal známým koncept percepčního centra, který byl rozpracován především v angličtině (Morton *et al.*, 1976): p-centra byla definována jako percepčně nápadné, periodicky se opakující události (kolem 500 ms), které se objevují poblíž začátku vokálu přízvukných slabik. P-centrum odpovídá místu ve slabice, které koresponduje s vnímanou rytmickou dobou. Ačkoli bývají p-centra spojována s vokalickými jádry, jejich umístění může variovat v souvislosti s fonotaktickou stavbou slabik, což bylo přesvědčivě prokázáno i v češtině (Šturm & Volín, 2016). Uspořádání p-center zřejmě ovlivňuje vnímání periodického rytmu nebo izochronní řeči (Fletcher, 2010). V experimentech v rámci této úrovně byly obvykle používány aktivity jako přizpůsobování stimulů vnímané izochronii (Marcus, 1981; Cooper *et al.*, 1986) nebo synchronizace řeči s metronomem (Barbosa *et al.*, 2005).

Izochronii na úrovni slabiky nebo stopy narušuje koncové nebo počáteční dloužení (Lehiste, 1977; Kohler, 1986) – taková nepravidelnost pak poskytuje vodítko k vnímání hranice pro posluchače. Ve výzkumu řečových chyb (Culter, 1981) se dále ukázalo, že přidání nebo vynechání slabiky v řeči nastává proto, aby vyrovnávalo počet slabik ve stopách a výpověď tak působila rytmičtěji.

Idea rozdělení jazyků do rytmických skupin podnítila řadu studií, které se snažily toto rozlišení vymezit konkrétněji pomocí poměřování **produkovaných** konsonantických a vokálních intervalů v předpokládaných taktově a slabičně izochronních jazycích (Grabe & Low, 2002; Ramus *et al.*, 1999). Přes všechnu tuto snahu nebyly jazyky do výše zmíněných skupin rozřazeny jednoznačně. Jak už předpokládal Pike (1945), některé jazyky vykazovaly v různé míře charakteristiky obou rytmických skupin – je pravděpodobné, že tyto rytmické vlastnosti jednotlivých jazyků tvoří oddělené skupiny, nýbrž spíše kontinuum. Nesporová (1990) ale navrhovala spíše než kontinuum vytvoření třetí, „prostřední“ kategorie pro jazyky, které zčásti splňovaly kritéria taktově izochronních a zčásti slabičně izochronních intervalů.

Ramus *et al.* (1999) představil jako relativně spolehlivé rytmické ukazatele co do typu izochronie  $\Delta C$ , směrodatnou odchylku trvání konsonantických intervalů, a %V, procentuální podíl vokálních intervalů v různojazyčných projevech. Na Ramuse navázal později Dellwo *et al.* (2007), když nahradil koreláty %V a  $\Delta C$  procentem znělých a směrodatnou odchylkou neznělých intervalů (taktově izochronní jazyk by měl mít nižší procento znělých a vyšší variabilitu neznělých intervalů), což je postup, který vykazuje spojitost s rozpoznáním rytmických skupin jazyků novorozenci (Ramus *et al.*, 1999). Podobně Lowová *et al.* (2001) dosáhla relativní přesnosti rozdělení jazyků do rytmických tříd za použití indexu párové variability, kterým je možné měřit variabilitu sekvenčního časování v následných intervokálních intervalech nebo přilehlých segmentových intervalech.

Kohler (2009) se však proti těmto přístupům vymezuje – zmíněné metody považuje pouze za způsoby třídění dat, které ale nevypovídají o pravé podstatě rytmu. Akustická analýza není v popisu rytmických jevů dostačující, jakkoli např. jevy spojené s trváním promluvy a trváním vokálů nebo směrodatnou odchylkou trvání konsonantů mohou napomáhat pochopení tradičního rozdělení jazyků do rytmických skupin. Podle Kohlera (2009) totiž tato měřítka nejsou schopna zachytit vzorce rytmických pohybů, které si

časem vyvíjejí pravidelnost. Jelikož je fenomén rytmu spíše než v objektivních vlastnostech signálu zakotven v percepci, je nutné při výzkumu rytmických charakteristik řeči zohlednit posluchače a přistupovat k výzkumu rytmických vlastností promluv komplexně – rytmus řeči je složitý jev, k jehož popisu nestačí pohled pouze z jediného úhlu.

Analýza rytmických aspektů řeči se v současnosti nemusí omezovat na trvání segmentů, slabik nebo stop. Pro produkci rytmu je důležité zejména globální temporální členění promluvy do částí, které disponují periodickými fonetickými charakteristikami. Tyto vlastnosti interagují se syntaktickou a sémantickou organizací, kterým jsou nadřazeny. Globální proměnné mohou být vzorce slabičného časování, základní frekvence, slabičné energie a spektrální dynamiky, které se opakují s určitou mírou pravidelnosti v čase (Kohler, 2009). Do centra pozornosti se přesouvá kromě posluchače i komunikativní funkce rytmu řeči (která se může projevovat například zjednodušením souhláskových shluků u slova v kontextu neplnovýznamovém a jejich pečlivější výslovností v kontextu plnovýznamovém), stejně tak se prosazuje tendence k upřednostňování přirozené řeči před laboratorní – je totiž možné očekávat, že výsledky výzkumů vedených na spontánních mluvených projevech budou více vypovídat o jevech v reálné komunikaci.

Zohlednění **komunikativního aspektu** při zkoumání rytmu řeči se projevilo již na konci 70. let a v 80. letech v experimentech s reakčními dobami. Bylo experimentálně prokázáno, že upravený rytmus zvýší u respondentů procento chybovosti; reakční doba byla také kratší u přirozeného řečového rytmu než u upraveného, byť percepčně přijatelného (Huggins, 1979; Buxton, 1983). Podobně v jiném experimentu (Ghitza & Greenberg, 2009) se zrychlenou přirozenou řečí, do níž byly vloženy buď stejně, nebo různě dlouho trvající tiché pauzy (od 20 do 160 ms), bylo prokázáno, že v případě vkládání pravidelných tichých intervalů trvajících do 120 ms srozumitelnost výrazně vzrostla, na rozdíl od případů s různě dlouhými tichými vsuvkami. Řeč s prediktabilní rytmickou strukturou je tedy vnímána s mnohem menším úsilím.

Kohler (2009) navrhuje nový přístup k výzkumu rytmu, který zahrnuje všechny čtyři hlediska – **jazyk, mluvčího, posluchače i komunikativní funkci** – a v jehož centru stojí posluchač a funkce. Zahrnuje čtyři proměnné pro signalizování rytmičnosti:  $f_0$ , trvání slabik, slabičnou energii a spektrální dynamiku. Tyto parametry včleňuje do experimentální metody sestávající ze tří částí: Nejprve musí být prozkoumána



prominence, a tedy relativní příspěvek uvedených čtyř parametrů k percepci prominence a tvoření rytmických vzorců. Druhá část se zabývá rytmickým provedením předepsané i improvizované řeči trénovaných i netrénovaných mluvčích a percepčním hodnocením jejich projevů; třetí část metody je aplikace prvních dvou kroků na různé jazyky pro jejich srovnání.

Nové výzkumné paradigma počítá s variabilitou v produkci i percepci jako s východiskem pro stanovení pružných rytmických vzorců; nepředpokládá povrchovou izochronii a dává posluchači klíčovou roli v rozhodování, co tvoří rytmickou pravidelnost. Rytmus projevu může být hodnocen v souvislosti s mluvčím, stylem i komunikační situací; má důležitou funkci pro posluchače. Tím pádem nelze ani prohlásit, že rytmus je jakýsi pevný vzorec stejný pro určitou skupinu jazyků, ale je variabilní i uvnitř každého jazyka, nicméně jazykem je do jisté míry předurčen k využívání různých potenciálních rytmických parametrů podle toho, jak jsou v daném jazyce užívány a navzájem spojeny na jiných jazykových rovinách. Analýza strukturně různých jazyků paralelně podle stejné metodologie může odhalit jazykově specifické užívání fyzikálních rytmických parametrů v řečové komunikaci, a tedy poskytuje širší základ pro výzkum rytmu (Kohler, 2009).

## **2.5 O rytmu řeči v češtině**

Jak bylo zmíněno výše, rozdělení jazyků do skupin pod štítky „taktově izochronní“ a „slabičně izochronní“ nefunguje bezvýhradně. Jedním z příkladů jazyka, který není snadno zařaditelný ani do jedné ze skupin, je právě čeština. Ačkoli byla čeština tradičně zařazována do skupiny jazyků slabičně izochronních (Palková, 1994), výsledky novějších zkoumání ukázaly, že tento západoslovanský jazyk nese i rysy druhého rytmického typu. Duběda (2004) aplikoval na češtinu měřítko indexu párové variability a Ramusova měření  $\Delta C$  a  $\%V$  (Ramus *et al.*, 1999) a zjistil, že první ukazatel češtinu zařadí po bok japonštiny a katalánštiny a staví ji tak mezi jazyky považované za slabičně izochronní, zatímco druhý umístí češtinu mezi jazyky taktově izochronní. Výzkum Dankovičové a Dellwa (2007) na základě komplexnosti českých slabik také nepřinesl jednoznačné rozřešení otázky rytmické klasifikace češtiny – vše tedy nasvědčuje tomu, že jazyky mohou vykazovat znaky obou rytmických typů a liší se pouze v míře jednotlivých rysů, jak už ostatně předvídal Pike (1945).

### 2.5.1 Faktory tvořící rytmické charakteristiky češtiny

Rytmickou strukturu češtiny je vhodné zkoumat v souvislosti s faktory, které produkci a percepci rytmu mohou ovlivňovat, jako je trvání hlásek, stavba slabik, taktový přízvuk, členění řeči do prozodických celků, větná prominence, mluvní tempo a v neposlední řadě i mluvní styl. Vzhledem k předpokládanému vlivu uspořádání segmentů v souvislé řeči na vnímání jejího rytmu je také vhodné se zaměřit na kombinační možnosti českých hlásek – fonotaktiku češtiny.

#### 2.5.1.1 Trvání hlásek v češtině

Na fonologicky nejnižší úrovni jsou důležitým faktorem temporální vlastnosti samotných hlásek. S výzkumy zaměřenými na měření časových aspektů českých hlásek začal již Chlumský (1911, 1928) a později na tuto větev navázaly studie v šedesátých letech (např. Kaiser, 1964; Borovičková & Maláč, 1967). Pro tyto starší studie jsou typické laboratorní podmínky získávání nahrávek – jednalo o časové hodnoty naměřené při čtení izolovaných vět či slov, v případě studie Borovičkové a Maláče (1967) dokonce izolovaných logatomů. Míra nepřirozenosti řeči pochopitelně mohla ovlivnit výsledky měření, jak je vidět na hodnotách ve studii Borovičkové a Maláče (1967), kteří v případě průměrného trvání vokálů naměřili u některých samohlásek více než dvojnásobné hodnoty než např. Psutka *et al.* (2006), který vycházel z korpusu spontánní řeči. Podobnou tendenci lze vysledovat i v publikaci Churaňové *et al.* (2015), která využívala paradigmatu synchronizace řeči s metronomem – jelikož toto paradigma pracuje s laboratorními podmínkami, naměřené temporální hodnoty českých hlásek byly přibližně 1,4× vyšší než v práci Weingartové (2015), jež k analýzám používala přirozenější materiál. V publikaci *Mluvíme s počítačem česky* (Psutka *et al.*, 2006) poskytli autoři komplexní data o trvání českých hlásek, temporální hodnoty však byly získány pomocí automatického přiřazování hlásek ke zvukovému úseku, což vedlo k vysokým odchylkám v časových hodnotách. Poslední prací, která zevrubně popsala trvání českých hlásek, je disertační práce Weingartové (2015). Pro tento popis byly využity čtené dialogy od 34 mluvčích. Plynulá řeč byla sice původně nasegmentována na hlásky automatickým algoritmem, ale později byly hranice hlásek kvůli přesnosti výsledků manuálně opraveny. Autorka při zpracování temporálních hodnot hlásek eliminovala faktory, které mohly průměrné trvání hlásek ovlivnit – pro co největší objektivitu tedy odstranila z analýz hlásky v pozici v závěrečném taktu v promluveném

úseku, v pozici iniciální hlásky po pauze či hlásky ze slov tvořících samostatnou prozodickou frázi; taktéž do výpočtu průměrného trvání českých hlásek nezahrnula výplňková slova spojená s hezitacemi, jako jsou *hm*, *no* a *jo* (Weingartová, 2015: 65, 111). Naměřené hodnoty z práce Weingartové (2015) jsou shrnuty v následujících tabulkách:

Vokály					
krátké	průměr	dlouhé	průměr	diftongy	průměr
<b>a</b>	58	<b>a:</b>	93	<b>ou</b>	80,3
<b>ε</b>	49,9	<b>ε:</b>	99,8		
<b>i</b>	45,6	<b>i:</b>	53,8		
<b>o</b>	47	<b>u:</b>	60,7		
<b>u</b>	45				

Tabulka 1. Reprezentativní průměrné trvání českých vokálů v milisekundách. Souhrn z Weingartová, (2015): 66. Trvání dlouhého ε: je pravděpodobně ovlivněno tím, že přes 50 % výskytů tohoto vokálu tvořilo slovo „ne“, kde dochází často ke dloužení (Weingartová, 2015: 114).

Konsonanty											
okluzivy	průměr	frikativy	průměr	afrikáty	průměr	nazály	průměr	vibranty	průměr	aprox.	průměr
<b>p</b>	78,1	<b>f</b>	59,6	<b>ts</b>	87,9	<b>m</b>	58	<b>r</b>	41,5	<b>l</b>	37,7
<b>b</b>	58,7	<b>v</b>	41,6	<b>dz</b>	66,7	<b>ɱ</b>	83,1	<b>r slb.</b>	71,1	<b>j</b>	28,6
<b>t</b>	68,1	<b>s</b>	92,7	<b>tʃ</b>	101,5	<b>n</b>	43,4	<b>ʀ/ ʁ</b>	53,8		
<b>d</b>	30,7	<b>z</b>	52,9	<b>dʒ</b>	63,7	<b>ɲ</b>	50,5				
<b>c</b>	71,9	<b>ʃ</b>	71,8			<b>ŋ</b>	86,9				
<b>ʈ</b>	52,1	<b>ʒ</b>	54,9								
<b>k</b>	64,4	<b>x</b>	65,8								
<b>g</b>	43,4	<b>h</b>	55,4								

Tabulka 2. Reprezentativní průměrné trvání českých konsonantů v milisekundách. Souhrn z Weingartová (2015): 112.

Jelikož hlavní složka výzkumné části prezentované disertační práce bude spočívat v percepčním testu s položkami původně proslovenými v rámci čtených textů, lze předpokládat, že hodnoty trvání českých hlásek naměřené Weingartovou (2015) budou aktuální i pro tento výzkum.

Uvedená trvání českých hlásek jsou pouze reprezentativními hodnotami. V aktuální řeči, ať už čtené, nebo spontánní, může trvání hlásek silně variovat. Faktory, které ovlivňují trvání hlásek, mohou být zároveň i činiteli působícími na tvorbu a vnímání rytmu řeči jako celku. Samostatné hlásky vykazují inherentní trvání, založené na fyziologických principech či charakteristikách specifických pro jazyk. Jedním z takovýchto principů je

fakt, že např. znělé šumové hlásky jsou v trvání kratší než jejich neznělé protějšky; stejný vztah obvykle platí pro vysoké a nízké vokály (Klatt, 1976; Beckman, 1986; Solé & Ohala, 2010) – tento fakt ověřila pro češtinu také Weingartová (2015). Kromě těchto inherentních faktorů je nutné počítat i s externími činiteli, jako je například struktura slabiky, artikulační tempo či prominence a pozice hlásky ve vyšší prozodické jednotce (Beckman, 1986; van Santen, 1992). I v češtině se ukazuje, že může hrát roli sémantický status slova – výsledky Volína a Weingartové (2014) naznačují, že plnovýznamová slova mohou být vyslovována pomaleji než neplnovýznamová. Na temporální vlastnosti samohlásek i souhlásek také může mít vliv mluvní styl (Chen, 1980; Picheny *et al.*, 1986).

Mluvní či artikulační tempo je jedním z nejzřetelnějších faktorů ovlivňujících trvání hlásek. Obecně lze konstatovat, že vyšší artikulační tempo má automaticky za následek zkracování hlásek (Klatt, 1976). Tento proces se ale neděje lineárně, některé segmenty jsou změnami tempa ovlivněny více než jiné. Pro češtinu Kaiserová (1964) zjistila, že krácení trvání hlásek je vyšší u fonologicky dlouhých vokálů než u krátkých. Rezistenci trvání krátkých vokálů oproti dlouhým při změně artikulačního tempa prokázala i Churaňová *et al.* (2015). Další výsledky přidal Machač (2006), který potvrdil temporální stabilitu českých exploziv. Jako významný prediktor temporálních změn hlásek v souvislosti s mluvním tempem se podle výzkumu Churaňové *et al.* (2015) ukázala třída hlásek – největší rozdíl v trvání hlásek ve dvou tempech vykazovaly dlouhé vokály, nazály a sibilanty, zatímco laterála [l] se jevila jako relativně stabilní v obou tempech. Tempo řeči ale v tomto výzkumu nemělo vliv na trvání konsonantů v souhláskových shlucích, což mohlo být dáno fyziologickými omezeními. Oproti tomu pozice segmentu v rámci vyšší prozodické jednotky (mluvního taktu) významně přispěla ke změnám trvání ve dvou tempech, a to zejména ve formě finálního dlužení slova: rozdíl v trvání hlásek mezi dvěma tempy byl větší ve finální pozici než v ostatních (Churaňová *et al.*, 2015).

Trvání segmentů může být také ovlivněno manifestací prominence na slovní nebo frázové úrovni. V řadě jazyků platí, že přízvukné slabiky inklinují k delšímu trvání segmentů než nepřízvukné (Crystal & House, 1988; Heldner & Strangert, 2001; Rietveld *et al.*, 2004; Sluijter & van Heuven, 1996a; Sluijter & van Heuven, 1996b). V češtině s pevným slovním přízvukem na první slabice mluvního taktu se ale prominence prodloužením hlásek neprojevuje (Chlumský, 1928; Janota & Palková, 1974). Nedávná

studie ale prokázala prodlužování segmentů na prominentních slabikách na úrovni větné – slova s frázovým (kontrastivním) přízvukem byla delší než ostatní (Volín, 2009).

Na lokální temporální charakteristiky má také vliv pozice hlásky ve vyšší prozodické jednotce (mluvní takt, promluvový úsek). Některé z těchto jevů jsou považovány za univerzální, například finální dloužení (Klatt, 1976). Přesný rozsah těchto jevů ale může být jazykově specifický i specifický pro mluvčího. Dankovičová (1997) konstatuje, že koncové dloužení se v češtině projevuje na úrovni promluvových úseků, ale míra dloužení může vykazovat variabilitu mezi jednotlivými mluvčími. Weingartová (2015) zjistila významné dloužení vokálů ve finálních slovech ve frázi – oproti vokálům v mediálních slovech jsou 1,5× delší, přičemž nejmenší rozdíl je u [u:] a největší u [u]. Rovněž expozivy mají tendenci se ve finálním slově prodlužovat, ale existují rozdíly v konkrétních segmentech: zatímco [t] se prodlužuje na 1,5násobek svého trvání v mediálních slovech, [p], se neprodukuje téměř vůbec. Frikativy se ve finálních pozicích také dlouží, přičemž neznělé více než znělé (Weingartová, 2015: 84). U neznělých afrikátů autorka konstatuje opět tendenci k finálnímu dloužení, avšak [tʃ] se prodlužuje výrazně více než [tʃ] (1,6násobek oproti 1,1násobku trvání v mediálních slovech). Nazály jsou také významně delší ve finálních slovech; pozice slova ve frázi má největší vliv na hlásku [m]. To se spolu s hláskou [n] prodlužuje nejvíce, na rozdíl od [n], které zůstává srovnatelně dlouhé v iniciálních, mediálních i finálních slovech. Efekt pozice slova je podle Weingartové (2015) významný i pro vibranty a aproximanty: hlásky [r], [ɾ] a [j] jsou nejkratší v mediálních slovech a nejdelší ve finálních, [l] je v iniciálních a mediálních slovech přibližně stejně dlouhé a zřetelně se prodlužuje až ve finálních. Nejvíce se prodlužuje právě [j], které ve finálních slovech může dosáhnout až 1,75násobku trvání v mediálních slovech, nejméně [r], které se prodlužuje maximálně na 1,1násobek svého mediálního trvání.

Variabilitu v trvání hlásek může způsobovat i fonetické okolí daných segmentů. Roli mohou hrát například fyziologická omezení jako vzdálenost artikulačních cílů jednotlivých fonémů či možnost překryvu artikulačních gest (Klatt, 1976). Klatt (1973) zjistil, že konsonanty ve shlucích bývají obecně kratší než v jednoduchých slabikách se strukturou konsonant-vokál (CV), ale tento vztah nebyl lineární pro všechny hláskové skupiny. Pro češtinu problematiku zpracovala Weingartová (2015), která zjistila, že vokály po rázu bývají kratší než v případech, kde se ráz nevyskytuje, a nejvíce se po rázu

zkracuje [u] (takřka na polovinu svého trvání) a nejméně [ɪ], které po rázu přichází pouze o 20 % svého trvání. Na trvání vokálů působí i typ slabiky – ke zkracování krátkých vokálů v jádru slabiky docházelo v materiálu Weingartové ve slabikách s préturou bez zjevné souvislosti s počtem konsonantů v prétuře. Z testovaných typů slabik byly typy V a VC signifikantně delší než ostatní typy CV, CVC, CCV, CCVC. Podle autorky na trvání vokálů neměla žádný vliv absence či přítomnost kody; výsledky Churaňové *et al.* (2015) ale naznačují statisticky významné dloužení vokálu ve finální slabice neobsahující kodu i rozdílné trvání finálních vokálů v souvislosti s typem kody (okluziva, frikativa, nazála). I na trvání expoziv působí hláskové okolí: intervokální okluzivy by měly být významně delší než okluzivy v souhláskové skupině v rámci slabiky i mimo ni. Jedinou expozivou, která se chová jinak, a je tedy mezi vokály naopak kratší, je [d] – u ní zřejmě dochází k redukci na alveolární švih, čímž se zkracuje i její trvání (Weingartová, 2015). Podobný trend jako u expoziv se potvrdil i u frikativ a afrikát. Nazály se ukázaly být také v intervokální pozici nejdelší a ve shluku nejkratší; jejich trvání v ostatních pozicích je delší než ve shluku a kratší než mezi samohláskami. U hlásky [j] byly rozdíly mezi pozicí ve shluku, mezi vokály a v ostatních pozicích nesignifikantní; hlásky [r] a [ɾ] byly kratší v intervokálních pozicích než ve shlucích a hláska [l] byla podobně krátká ve shluku i mimo něj; naopak se prodlužovala v ostatních pozicích.

Weingartová (2015) také pozorovala určité – relativně různorodé – tendence k vlivu délky prozodické fráze ve slabikách i slovech na trvání hlásek. Jisté trendy ukázal i výzkum Churaňové *et al.* (2015), který mimo jiné porovnával trvání iniciálních a finálních konsonantů v jedno- až trojslabičných slovech a potvrdil hypotézu, že segmenty inklinují ke kratšímu trvání ve víceslabičných prozodických jednotkách. Problematika by si nicméně zasloužila podrobnější přezkoumání.

#### **2.5.1.2 Slabika a výzkum p-center v češtině**

Základním principem souvislé lidské řeči je střídání apertury a striktury. Tento kontrast se promítá do struktury elementární jednotky souvislé řeči, slabiky – její nejobvyklejší varianta se skládá z kombinace konsonantu v prétuře a vokálního jádra (CV). *Teorie rámce a obsahu* (MacNeilage, 1998) uvádí, že ústní komunikace se vyvíjela původně z pohybů čelisti typických pro mechanismus příjmu potravy – sekvence otevírání a zavírání čelisti tedy vytvořily řečový *rámec* a řízení motoriky ostatních artikulátorů se

později vyvinulo ke globální koordinaci pro generování *obsahu* (Rochet & Schwartz, 2005). V mnohých jazycích včetně češtiny se však postupně vyvinuly rozmanité typy slabičných struktur. Roussetová (2004) zkoumala zastoupení různých slabičných typů v 16 jazycích a dospěla k závěru, že nejčastějšími slabičnými strukturami jsou typy CV a CVC, méně obvyklý výskyt vykazovaly typy V, CCV a VC, CCVC a CVCC. Ostatní struktury nedosahovaly ani jednoho procenta ve vzorku.

Zatímco v préture české slabiky se mohou objevit čtyři konsonanty a ve výjimečných případech pět (např. ve slovech *vzkvět*, *k pštrovi*), v kodě mohou být souhlásky nejvýše tři. Tento fakt je v souladu s univerzální tendencí k maximálnímu zaplnění préture (např. Blevins, 1995). Z těchto skutečností je množné usoudit, že v českých slovech a taktech často vznikají souhláskové shluky, nicméně skupiny tvořené více než třemi konsonanty se v projevech objevují vzácně (Volín & Churaňová, 2010). Je také nutné zmínit, že pro češtinu nejsou stanovena jednoznačná pravidla pro určení hranic slabiky, a proto mohou být poznatky o četnosti výskytu jednotlivých slabik a slabičných typů poněkud zkresleny – např. dvouslabičné slovo *sestra* může být interpretováno jako sled slabik *ses-tra* (CVC-CCV), *sest-ra* (CVCC-CV) i jako *se-stra* (CV-CCCV).

Struktura slabiky je důležitým faktorem percepce rytmu jazyka. Vztah mezi slabičnou strukturou a rytmem mluvené češtiny byl zkoumán v nedávných studiích Volín *et al.* (2014), Churaňová (2015) a Šturm & Volín (2016) s pomocí paradigmatu synchronizace řeči s metronomickými pulzy, které v těchto případech suplovaly percepční (či produkční) centrum první slabiky. Pro účely výzkumu byla použita dvouslabičná existující česká slova s kontrolovanou fonotaktickou strukturou. Préture první slabiky se skládala z jednoho, dvou, nebo tří konsonantů; první slabika neměla kodu. Préture i koda druhé slabiky byla reprezentována jedním konsonantem.

24 testovaných subjektů slyšelo pravidelné metronomické pulzy a opakovaně vyslovovalo cílové slovo zobrazené na monitoru. Realizace slova mluvčí synchronizovali s izochronními pulzy tak, aby každá slyšitelná doba procházela první slabikou každého izolovaného slova. Mluvčí se měli snažit o co nejpřirozenější realizaci daných slov. Každá položka obsahovala 12 pulzů a desenzitační pasáž; mluvčí začal vyslovovat slovo s pátou dobou. První čtyři pulzy sloužily jako příprava pro mluvčího, který se tímto mohl snadněji vpravit do požadovaného rytmu. Subjekt vyslovil slovo celkem osmkrát; první a poslední realizace byly z dalších analýz vypuštěny. Všechny položky byly prezentovány

ve dvou různých tempech – „normálním“ a „rychlém“. V prvním případě se pulz vyskytoval s frekvencí 70 dob za minutu (každých 857 ms), ve druhém pak s frekvencí 90 úderů za minutu (každých 667 ms). Tyto frekvence byly vybrány, aby podněcovaly produkci mluvních temp cca 4 a 6 slabik za sekundu.

Výsledky experimentu ukázaly, že synchronizace prvního vokálu v daném slově s metronomovým pulzem se měnila v souvislosti s komplexností préture první slabiky. Složitě préture způsobovaly, že se pulz vyskytoval dále před začátkem vokálu; podobný jev vyvolávala i přítomnost dlouhého vokálu v první, přízvučné slabice. Slabší efekt na umístění pulzu v první slabice slov byl vypořádan i v případě přítomnosti a typu kody slabiky druhé. Výsledky ukázaly, že nulová koda posouvá pulz blíže začátku vokálu, zatímco u slov, jejichž druhá slabika kodu obsahovala, byla vysledována tendence k posunu pulzu dále před začátek vokálu (Volín *et al.*, 2014). Určitá tendence byla pozorována, i co se týče typu kody: zatímco obstruent v kodě byl spjat s nárůstem synchronizačního intervalu o 18 ms v pomalém tempu a o 11 ms v tempu rychlém, sonorní koda vedla k prodloužení intervalu synchronizace pouze o 8 ms v pomalém tempu a o 15 ms v tempu rychlém – tempo tedy ovlivňovalo obstruenty v kodě opačným způsobem než sonory na témže místě (Šturm & Volín, 2016).

V některých jazycích byly také pozorovány temporální kompenzace mezi vokálem a sousedním konsonantem. Obvykle samohlásky inklinují ke kratšímu trvání, pokud následují neznělé souhlásky, které zpravidla trváním znělé převyšují (Klatt, 1976; Lisker, 1974; Port *et al.*, 1980; Kent & Read, 2002). Podobný jev, avšak v opačném směru, zjistili Borovičková a Maláč (1967) pro češtinu – vokál následující po neznělé souhlásce měl kratší trvání než po znělé. V češtině zatím kompenzační strategie nebyly systematicky zkoumány, nicméně nedávné výzkumy temporálních interakcí mezi segmenty ukázaly na možnou kompenzaci mezi výškou vokálu a předcházející explozivou – ta inklinuje ke kratšímu trvání, pokud je následující samohláska nízká (Machač, 2006). Studenovský (2012) zjistil kompenzační chování mezi diftongem a počtem segmentů v préture slabiky – pokud byly v préture konsonanty dva, diftong byl kratší než po jednoduché préture.

### **2.5.1.3 Taktový přízvuk v češtině**

V češtině je taktový přízvuk pevně svázan s první slabikou mluvního taktu: netvoří ani nemění význam slova. V souvislosti s přízvukem byl často zmiňován obecný



(a z dnešního hlediska poněkud neprůhledný) pojem „síla“. Hála (1948: 142–143) konstatuje, že „základním pojítkem taktu je síla“ a že výpověď je členěna pomocí přízvuku a důrazu, které jsou oba „projevem síly“. V práci z roku 1962 už je Hála konkrétnější: uvádí, že podstatou přízvuku v češtině je zesílení hlasu při artikulaci přízvučné slabiky, což by mělo vést ke zvýšení sonority vokálů a k „zesílení“ a prodloužení konsonantů v této slabice, kteréžto jevy spolu s možným zvýšením základního tónu přispívají k prominenci přízvučné slabiky (Hála, 1962: 303). Autor ovšem předchozí popis přízvučnosti poněkud zmírňuje tvrzením, že sílu přízvuku není možné považovat za absolutní hodnotu – přízvučná slabika v celkově tiché řeči nemusí dosahovat síly nepřízvučné slabiky v řeči hlasitější. Proto je třeba přízvuk hodnotit srovnáváním slabik v rámci jednoho slova nebo taktu (Hála, 1975). Český přízvuk navíc autor považuje za „mírný“ a rozdíl v síle přízvučné a nepřízvučné slabiky za „nevelký“ (Hála, 1962: 303). Podobný náhled na přízvučnost prezentuje i Ondráčková (1954), která také považuje přízvučnou slabiku v rámci taktu vždy za „nejsilnější“. Dalším autorem, který operuje s termínem „síla“ v souvislosti s přízvukem, je František Daneš. Ve své práci z roku 1957 uvádí, že přízvuk pravděpodobně tvoří různý stupeň síly jednotlivých vyslovovaných slabik a hlásek. Autor také zpochybňuje názor, že je „silový přízvuk“ založen na intenzitě, a uvádí, že za jev známý jako slovní přízvuk může ve skutečnosti odpovídat několik faktorů, a ne pouze jeden. Také se podobně jako Hála (1975) a Ondráčková (1954) přiklání k názoru, že v případě taktového přízvuku nejde o absolutní hodnotu „síly“, nýbrž o „silový poměr“ jednotlivých slabik ve vyšší prozodické jednotce (Daneš, 1957: 21–22). V čem se tyto práce shodují, je fakt, že český přízvuk není záležitostí absolutní, nýbrž relativní prominence přízvučné slabiky vůči ostatním nepřízvučným slabikám v rámci taktu nebo vyšší jednotky. Ondráčková (1962) v analýze přízvučnosti dochází k podobnému závěru, totiž že český přízvuk je komplexní jev, jehož podstatou je spíše zvukový kontrast, který rozlišuje přízvučnou slabiku od nepřízvučných, než jednoznačná, akusticky vymežitelná vlastnost. Janota a Palková (1974) zjistili, že v některých pozicích může být přízvučná slabika dynamicky slabší a mít kratší trvání než slabiky okolní (Janota & Palková, 1974: 58). Autoři dále konstatují, že percepční hodnocení předělů taktů (a tedy i vyhodnocení určité slabiky jako přízvučné) je vhodné analyzovat na materiálu přirozené řeči a v širším řečovém kontextu, v němž se spíše mohou projevit interakce faktorů, které mají na vnímání přízvučnosti v češtině vliv.

Z analýz českého přízvuku tedy vyplývá, že nebyla jednoznačně definována určitá akustická charakteristika, která by slovní přízvuk systematicky vyznačovala (Palková, 2004). Přízvukná slabika českého mluvního taktu nemívá stabilně delší trvání, vyšší frekvenci ani intenzitu a samohlásky v ní mají stejnou kvalitu jako vokály ve slabikách nepřízvukných. Jednotka mluvního taktu tedy zřejmě není opřena o akustické kvality přízvukné slabiky. Vnímání taktu jakožto ucelené jednotky je zřejmě způsobeno průběhem zvukových charakteristik v sekvenci slabik, díky němuž vzniká dojem samostatnosti prozodické jednotky (Palková *et al.*, 2004).

V některých starších pracích (Havránek & Jedlička, 1988; Daneš, 1957; Hála, 1962) figuruje tzv. *vedlejší přízvuk*, který by se měl vyskytovat převážně na lichých slabikách ve slovech delších než dvouslabičných a který by měl být slabší než přízvuk na prvních slabice. V běžném projevu se ale s takovýmto jevem příliš často nesetkáme, je typický zejména pro slabikování či skandování – a i v těchto případech nastává otázka, jestli přítomnost dalšího přízvuku ve slově spíše nepovažovat za znak rozpadu slova na dva mluvní takty.

#### **2.5.1.4 Prozodická hierarchie souvislé češtiny**

Nejnižší jednotkou prozodické struktury češtiny je slabika. Za vyšší jednotku je považován tzv. *mluvní takt*, který může obsahovat jednu a více slabik. Na dalším stupni prozodické hierarchie češtiny tradičně stojí promluvový úsek, složený z jednoho či několika taktů; nejvyšší úroveň pak zastupuje výpověď neboli promluva – jeden či více promluvových úseků, typicky tvořících větu. Tato třístupňová hierarchie je přítomna v mnoha prozodických konceptech (Hirst & Di Cristo, 1998: 36) založených na různých jazycích – lze tedy předpokládat, že v základním prozodickém členění můžeme najít rysy, které jsou pro jazyky univerzální.

##### **2.5.1.4.1 Mluvní takt**

Takt je považován za nejmenší jednotku, na níž se realizuje rytmus řeči a je rovněž základní rytmickou jednotkou pro členění souvislé mluvy. Definice mluvního taktu se v pojetí rozdílných českých autorů v průběhu 20. století různily (Ondráčková, 1954 a Churaňová, 2012); prominentní však zůstávala představa delimitace taktu přízvuknou slabikou (Chlumský, 1928; Trávníček, 1932; Hála, 1948; Romportl, 1985).

Ačkoli byl český takt tradičně definován jako posluchačem vnímaná skupina slabik podřízená právě jednomu přízvuku (Ondráčková, 1954), z dalších výzkumů druhé poloviny 20. století a z nedávných studií vyšlo najevo, že uvažovaná stěžejní role přízvučné slabiky není v percepci posluchače ani zdaleka tak podstatná jako celkový průběh zvukových charakteristik v rámci této jednotky:

*„Mluvní takt lze v češtině charakterizovat jako lineární jednotku s vnitřní kohezí. Relevantní vlastností je zejména kontura F0. Podíl může mít i rozdíl v trvání segmentů sousedících jednotek.“* (Palková, 2004: 401)

Mluvní takt je tedy v percepci pravděpodobně vnímán zejména díky své intonační soudržnosti, nikoli na základě delimitace přízvučnou slabikou. Palková (2004) dále uvádí, že přízvučná slabika je tedy pouze odvozeným pojmem – jedná se o iniciální slabiku v taktu. Je proto vhodnější v terminologii obvyklý pojem „přízvukový takt“ nahradit taktem „mluvním“ (Palková *et al.*, 2004).

Taktu se na úrovni jazykové struktury blíží jednotka slova. Mezi slovo a takt však nelze položit rovnítko – mluvní takt může obsahovat jedno, ale i více slov; v přirozených podmínkách a zejména ve spontánním projevu se můžeme setkat i s případy, kdy je jedno slovo rozděleno do více taktů (Hála, 1948; Ondráčková, 1954; Churaňová, 2013).

Čeština jakožto jazyk s pevným přízvukem nevyužívá slovní prominenci k signalizaci významu slova. Přízvuk je však jedním ze zvukových rysů taktu – jak již bylo zmíněno výše, bývá přízvuk na první slabice taktu. Ani toto pravidlo však neplatí bezvýhradně – existují případy, kdy je přízvuk na slabice druhé a první slabika tak tvoří tzv. *taktovou předrážku*. Typickými taktovými předrážkami bývají neplnovýznamová slova; v materiálu Ondráčkové (1954) se nejčastějšími předrážkami ukázaly být spojky (*a, že, když, i*), zájmena (*co, to*), zájmenná příslovce (*jak*) nebo sponová slovesa (*je*). Ač někteří autoři (např. Hála & Sovák, 1947; Ondráčková, 1954) považovali taktové předrážky za samostatné takty, v současnosti je standardem pojetí reprezentované Danešem (1957), které předrážku započítává do taktu následujícího.

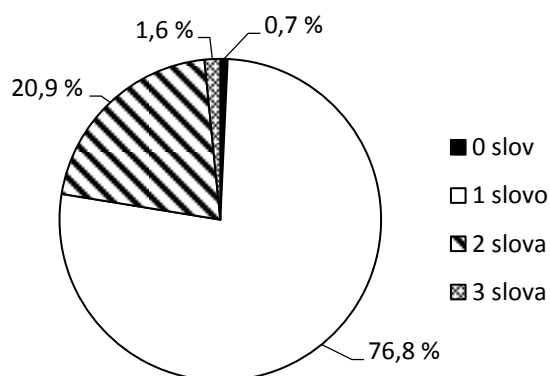
Ačkoli členění na takty v češtině není zcela ustálené, je možné zde vysledovat několik pravidelných tendencí: Víceslabičná slova tvořívají samostatné takty – ani v tomto případě však nejde o neporušitelné pravidlo; v přirozených podmínkách je relativně

běžné spojování dvou dvouslabičných slov nebo dvouslabičného a trojslabičného slova do jednoho taktu (Churaňová, 2012: 82):

/ 'Podle / 'lékařských / 'zpráv / '**není jeho** / 'život / 'bezprostředně / 'ohrožen. /

Jednoslabičná slova ve spontánním či připraveném projevu ani ve čtené řeči nemusí vždy stát v pozici taktové předrážky – např. pokud jde o původní předložku, nese tato obvykle přízvuk a stojí na začátku taktu (až na několik specifických případů, viz Churaňová, 2012); v případě, že se jedná o neplnovýznamová jednoslabičná slova, může být jejich postavení v proudu řeči různé – mohou přízvuk zachovat a tvořit samostatný takt, či stát na jeho začátku, ale mohou přízvuk i ztratit a přiklánět se k předcházejícímu, nebo následujícímu taktu. Možností rozčlenění textu se obvykle nabízí několik a záleží jen na mluvčím, jak bude text frázovat. V rámci jednoho taktu se navíc mohou jednotlivé tendence kombinovat, např. ke dvouslabičnému slovu se může připojit předcházející i následující jednoslabičné slovo (Churaňová, 2012: 83):

/ **Jak 'hlásí náš** / 'zpravodaj / 'v Německu / 'Jiří / 'Hošek / (...)



Graf 1. Procentuální zastoupení daného počtu slov v rámci taktů v českých textech. 0 slov = rozpad slova na dva takty. Převzato z Churaňová (2013).

počet slabik	počet výskytů	% ze všech taktů
2	1816	33,83
3	1699	31,65
4	1089	20,29
1	363	6,76
5	308	5,74
6	72	1,34
7	12	0,22
10	4	0,07
9	2	0,04
8	1	0,02
11	1	0,02
12	1	0,02

Tabulka 3. Počet slabik v taktech a přepočet na procenta (100 % = 5368). Převzato z Churaňová (2013).

V češtině se za nejčastější variantu taktu podle délky považuje takt dvouslabičný, po něm následuje tříslabičný (Hála, 1948; Ondráčková, 1954; Churaňová, 2013). Nezanedbatelné zastoupení co do frekvence v mluvených projevech mají i takty čtyřslabičné a jednoslabičné; ostatní takty už v češtině natolik obvyklé nejsou

(Ondráčková, 1954; Churaňová, 2013). Co se týče poměru slov a taktů, výzkum Churaňové (2013) prokázal, že nejčastější variantou taktu je takt jednoslovný; pětinu všech taktů v materiálu ale tvořily takty dvouslovné.

#### 2.5.1.4.2 Promluvvový úsek

Za prozodickou jednotkou vyšší než mluvní takt bývá označován promluvvový úsek. Ten je v češtině definován jako sled mluvních taktů vázaných intonačními prostředky či pauzami tak, že je posluchač vnímá jako jediný zvukový celek. Promluvvový úsek není výhradně percepční záležitostí – i sami mluvčí inklinují ke tvoření těchto celků. Palková (2004) ho označuje za základní zvukovou jednotku z hlediska komunikačních cílů mluvčího.

Popisem promluvvového úseku se zabýval Daneš (1957), který úsek vymezil jako jednotku určenou zejména hraničními signály, tvořenými specifickým intonačním průběhem či pauzou, a tzv. *intonačním centrem*. Intonační centrum Daneš považoval za intonační prominenci inherentní každému úseku a stojící zpravidla na konci jednotky. Palková (1994) s Danešovým pojetím polemizuje a konstatuje, že pokud bude intonační centrum podmínkou každého úseku, ztratíme možnost rozlišení zvukové a významové prominence, neboť k oběma typům prominence se v řeči používají stejné prostředky. Dále autorka uvádí, že v percepci posluchače není promluvvový úsek spojen s jedinou prominencí – některé úseky mohou být vnímány s prominencí na svém začátku i konci či bez jakékoli prominence.

Promluvvový úsek bývá vymezen zejména intonačním pohybem (zejména v závěru intonační jednotky), pauzou, finálním dloužením a dalšími jevy, jako je např. finální snížení intenzity, globální konfigurace prozodických parametrů přispívajících ke koherenci úseku, kontrast se sousedními přízvukovými jednotkami apod. (Duběda, 2005: 190–191).

Promluvvové úseky mohou co do délky značně variovat. Podle Palkové (1994) obsahují nejčastěji úseky 2–3 takty a 6–7 slabik. Tento počet taktů a slabik ve frázi se může zdát poněkud nízký – záleží ale na tom, zda rozlišujeme různé hloubky předělu v rámci úseků. Hranice některých úseků jsou vyznačeny pouze intonačním průběhem, jiné jsou ovšem určeny i finálním dloužením nebo pauzou či dalšími charakteristikami, a jsou tedy výraznější. Pokud bychom za promluvvový úsek považovali pouze část promluvy

ohraničenou předělý vyššího řádu, byl by průměrný počet slabik či taktů v úseku pravděpodobně vyšší, než kdybychom pracovali s každým úsekem vymezeným hranicí jak vyššího, tak nižšího řádu.

Ačkoli má promluvový úsek relativně blízký vztah k syntaktické struktuře věty, není možné úsek a frázi či syntagma považovat za ekvivalentní. Vliv na prozodickou organizaci má rytmický faktor, syntaktická struktura a také tendence vázat k sobě významové komplexy. Vztah textu a frázování je oboustranný: frázování může, ale nemusí odpovídat stavbě textu, stejně tak struktura textu může, ale nemusí frázování podporovat. Některé parametry také spíše podporují textovou kohezi, jiné naopak jeho členění. Důležitou roli hraje počet taktů (promluva složená z více než čtyř taktů podporuje členění promluvy; čtyři a méně taktů obvykle tvoří dohromady jeden úsek), syntaktická struktura výpovědi (nestejná úroveň těsnosti závislosti mezi takty podporuje členění, rovnocenně těsné vazby mívají za následek soudržnost textu; složitější výpověď může být realizována s dvoustupňovou hierarchií předělů), sémantická stavba výpovědi (sémantická nesamostatnost podporuje textovou kohezi) a rytmická struktura taktů v úseku (střídání rytmických schémat může podpořit členění v místě začátku opakování rytmické sekvence). Všechny tyto parametry také mohou vzájemně interagovat a jejich účinek může být rozdílně silný – izolace těchto parametrů proto není snadná (Palková, 2006).

#### *2.5.1.4.3 Promluva*

V deskripci české souvislé řeči se dále vyděluje promluva neboli výpověď – prozodická jednotka na úrovni věty. Promluva jakožto popisná jednotka se definuje jako uzavřený celek s charakteristickým zvukovým průběhem, zejména na svém konci – v pracích Romportla (1951, 1985), Daneše (1957) a Palkové (1994) jsou rozpracovány jednotlivé melodické typy, charakteristické pro jednotlivé druhy výpovědí (např. oznámení a otázka doplňovací, otázka zjišťovací, neukončená výpověď, emocionálně zabarvená výpověď). V promluvě může být jeden či několik promluvových úseků; hranice promluvového úseku se tedy vždy shoduje s hranicí promluvy. V promluvě i promluvovém úseku se může objevit jedna nebo více prominencí vyšší úrovně než roviny taktu – tzv. *větný* či *kontextový přízvuk*.

### 2.5.1.5 Větná prominence

Jednou z možností, jak zvýraznit slovo nebo frázi v promluvě, může být zdůraznění pomocí *větné prominence*. Počet a umístění větných prominencí jsou také faktory, které mohou ovlivnit vnímání rytmické výstavby výpovědi. Ani tento druh prominence není charakterizován jednoznačně konkrétními vlastnostmi – jako zdůraznění může posloužit použití příznakové melodické kadence na taktu, který chce mluvčí zdůraznit, ale také změna dynamiky (např. zvýšení hlasitosti) při proslovení daného slova či fráze.

Romportl (1985) spojuje *větný přízvuk* se slovem na konci věty, které zároveň tvoří jádro (réma) výpovědi v terminologii aktuálního větného členění. Přítomnost větného přízvuku podle něj souvisí již s přítomností melodické kadence, která bývá obvykle na posledním taktu v úseku či promluvě. Větná prominence ale může být i na jiných místech v úseku než na jeho konci, a stejně tak se k jeho realizaci nemusí použít pouze intonační či dynamické prostředky: při vytýkání se může uplatnit i změna artikulačního tempa, pečlivější výslovnost jednotlivých hlásek ve zdůrazněném slově, oddělení slova od ostatních pauzami apod. Tyto prostředky se mohou vyskytovat v kombinaci i jednotlivě, jejich úkolem je vytvořit v daném kontextu kontrast. Tento kontrast může dále přispět i k vnímání hranic promluvových úseků (Palková, 1994: 298–300).

Větný přízvuk se může objevit jako součást přirozené modulace řeči, typicky na konci úseku, nebo může být spjat s významem slova. Palková (1994) považuje oba tyto druhy prominence za fakultativní, zatímco Daneš (1957) považuje větný přízvuk (či v jeho terminologii „intonační centrum“) za povinnou součást každého úseku a každé realizace melodického schématu (Daneš, 1985). Zdůraznění slova či fráze je nutné pouze v případě, že vyjadřuje vztah mezi jednotlivými slovy a tento vztah zároveň není signalizován jinými prostředky, např. slovosledem. Pokud větná prominence slouží k vyjádření vztahu mezi konstituenty textu, nazývá se *kontextový přízvuk* (Palková, 1994). Slovo zdůrazněné kontextovým přízvukem bývá takové, které se již v předcházejícím kontextu objevilo.

#### 1. *Budu vařit "brambory."*

Větná prominence na posledním slově promluvového úseku je fakultativní a může být výsledkem automatické modulace hlasu. Pokud by mluvčí chtěl zdůraznit, že hodlá vařit

brambory, a nikoli rýži, a za tím účelem použít kontextový přízvuk, musí na poslední slovo položit velmi výraznou realizaci přízvuku.

*2a. Budu "vařit brambory.*

Kontextový přízvuk na druhém slově upřesňuje typ tepelné úpravy brambor – nehodlá brambory smažit nebo péct.

*2b. Brambory budu (")vařit.*

Tato věta vyjadřuje stejnou skutečnost jako věta 2a. Použití kontextového přízvuku je však fakultativní – informace je sdělována slovosledem.

*3a. "Budu vařit brambory.*

Kontextový přízvuk na prvním slově má funkci ujištění, že mluvčí hodlá vařit, pokud má o tom posluchač pochybnosti.

*3b. Brambory vařit (")budu.*

Tato věta vyjadřuje stejnou skutečnost jako věta 3a. Použití kontextového přízvuku je však fakultativní – informace je sdělována slovosledem.

### **2.5.1.6 Mluvní tempo a pauzy**

#### **2.5.1.6.1 Mluvní tempo**

Mluvní a artikulační tempo tvoří další významný faktor při vytváření i percepci rytmu promluvy. *Mluvním tempem celkovým* se obvykle rozumí tempo řeči celého projevu určené jednotkami řeči za jednotku času (zpravidla slabikami za sekundu); *tempo artikulační* zahrnuje pouze řeč s vynecháním všech pauz kromě závěrových fází okluziv (Koreman, 2006). Pro účely výzkumu se používá také *mluvní tempo modifikované*, v němž se pro analýzu vynechávají pauzy určitého typu nebo nad určitý časový limit podle potřeb aktuálního experimentu (Palková *et al.*, 2004).

Podobně jako ostatní ukazatele nelze vnímání tempa řeči považovat za jednoznačný ukazatel – subjektivní tempo totiž nemusí korespondovat s tempem objektivním. Například i objektivně relativně rychlé tempo (cca šest slabik za sekundu) může být posluchačem vnímáno jako „středně rychlé“ (Veroňková & Janoušková, 2010).

Mluvní tempo je charakteristikou závisející na několika dalších faktorech – tempo se totiž může lišit v souvislosti s řadou dalších faktorů. Balkó (2003) dělí tyto faktory na tři



skupiny – individuální, nadindividuální a jazykové/sociokulturní. Mezi individuální faktory řadí vše, co přímo souvisí s osobností mluvčího: jeho emocionalitu a temperament, psychický a fyzický stav, pohlaví, věk a jazykovou kompetenci; faktory nadindividuální jsou reprezentovány stylem komunikace, prostředím, funkcí, tématem a typem projevu dle připravenosti (např. připravený × spontánní); poslední skupina zahrnuje jazyk, kterým je projev prosloven, a „národní temperament“.

Ačkoli se nedá jednoznačně určit, zda je v češtině obecně „rychlejší“ čtený, nebo spontánní projev, odborníci se shodují na tom, že ve spontánních či semispontánních projevech bývá co do tempa větší variabilita (Zima, 1959; Bartošek, 2000), ačkoli např. Rubovičová (2014: 50) došla k odlišnému závěru. Větší variabilitu tempa zaznamenal Zima (1959) také u projevů dialogických, které v jeho výzkumu střídaly různá tempa, oproti relativně stabilním projevům monologickým.

Mluvní tempo části projevu může také souviset s prozodickou jednotkou a pozici měřeného úseku v textu. Dankovičová (1997) prokázala, že artikulační tempo silně variuje v rámci promluвовého úseku. Balkó (2001) zkoumala tempo částí zpráv o počasí a dospěla k závěru, že tempo řeči má tendenci být vyšší na začátku a konci projevu, zatímco ve středu textu inklinuje ke snížení. Tuto tendenci potvrdila ve svém experimentu i Rubovičová (2014). Výsledky se mohou zdát v kontradikci s poznatkem o finálním dloužení, tj. zpomalováním tempa řeči na konci vyšší prozodické jednotky – aby se totiž tomuto efektu autorka vyhnula, podrobila analýze věty, které samotnému konci projevu předcházely, a tudíž se na nich efekt závěrového zpomalování neprojevil.

Dankovičová (1998: 72nn) našla silný efekt přizpůsobování se artikulačního tempa délce mluvního taktu co do počtu slabik: její výsledky ukázaly na přímou úměru mezi počtem slabik v taktu a rychlostí artikulačního tempa. Tento fakt ukazuje na možnost temporální kompenzace. K podobnému závěru došla i Balkó (2005) na prozodické rovině o jeden stupeň vyšší: čím více taktů v nahrávkách jejího výzkumu tvořilo promluвовý úsek, tím bylo kratší průměrné trvání taktu v úseku.

Na změny tempa v rámci projevu působí i textový či gramatický faktor. Romportl (1958) například v této souvislosti zmiňoval parenteze, které byly prosloveny jiným (nižším i vyšším) tempem než jejich okolí.

Mluvní či artikulační tempo, které je v češtině považováno za neutrální, se v průběhu 20. století v názorech odborníků měnilo. Studie většinou uvádějí, že se v průběhu času čeština „zrychluje“ (Bartošek, 2000; Veroňková, 2012). V současnosti je za neutrální považováno průměrné tempo pěti slabik za sekundu (Palková, 1994; Krčmová, 2005).

#### *2.5.1.6.2 Pauzy*

Duběda (2005) definuje pauzu jednoduše jako projev diskontinuity řeči a rozděluje pauzy dle jejich typu na pauzy fyziologické (nádechové), jazykově strukturní (syntaktické a sémantické) a akcidentální (hezitační). Müllerová a Nekvapil (1986) pauzu vymezují z hlediska recipienta – jako přerušení proudu řeči u mluvčího identifikovatelné posluchačem. Do své klasifikace nezahrnují pauzy nádechové, ale přidávají do ní pauzy kontaktové, které se objevují za tázacími dovětky a otázkami v rámci dialogu. Pauzy lze také klasifikovat co do vnímaného trvání (Romportl, 1958) či jejich obsahu (vyplněné či tiché, např. Rubovičová, 2014).

Pauzy rovněž úzce souvisejí s vnímáním rytmu – podle studie Zimy (1959) zkrácení pauz o cca polovinu jejich trvání vede k vnímání řečového stimulu jako prosloveného rychlejším tempem. Sedláková (později Dankovičová, 1989) rozlišuje pauzy na rušivé a nerušivé z hlediska posluchače. Pauzy nerušivé, logické, mají přiměřené trvání a jsou umístěny na správném místě z hlediska stavby textu; pauzy rušivé mohou být buď tematické, které přeruší téma hovoru, nebo nelogické – nesprávně umístěné z hlediska textové výstavby, či umístěné správně, ale s příliš dlouhým trváním. Také podle Balkó (1999) mají pauzy vliv na hodnocení projevu posluchačem – podle autorky jsou příjemněji vnímány projevy s více pauzami, které jsou v projevu logicky umístěny. Posluchači totiž takové pauzy usnadňují porozumění textu, protože naznačují jeho strukturu, a také dávají adresátovi projevu čas na zpracování přijímané informace.

Pauzy jsou důležité také pro mluvčího – poskytují mu totiž čas na rozmyšlení pokračování textu, na výběr lexikálních jednotek a na jejich zpětnou reflexi, popř. opravu (Bartošová, 2016).

#### **2.5.1.7 Mluvní styl**

Jak již bylo zmíněno v předchozí podkapitole, temporální vlastnosti projevu může ovlivnit styl, jakým je řeč proslovena. Větší variabilita tempa se v češtině vyskytuje spíše v projevech spontánního či semispontánního projevu (Zima, 1959; Bartošek, 2000)

a také v projevu dialogickém oproti monologickému (Zima, 1959). Trvání jednotlivých jednoslabičných slov v různých mluvních stylech porovnávala Ondrušková (2011) a dospěla k závěru, že trvání slov je variabilnější ve spontánních projevech než v projevech čtených, což vysvětluje několika faktory, např. vyšší formálností u čteného projevu, a tedy i pečlivější výslovností ve srovnání se spontánním projevem, kdy dochází k elizi hlásek. Autorka také zmiňuje problematiku dialogických projevů – zatímco v monolozích není mluvčí přerušován komunikačním partnerem a nemusí se mu přizpůsobovat, v dialogu musí mluvčí na partnera reagovat, dávat najevo, že hodlá repliku ukončit (poklesem intonace), nebo naopak v projevu pokračovat (prodloužení vokálu v poslední slabice slova před předělem).

Ačkoli pro češtinu neexistuje příliš srovnatelných dat pro vyslovení jednoznačného závěru ohledně vlivu konkrétního stylu na rytmus řeči, určitou představu si můžeme udělat z výzkumu Bartoška (2000), který srovnával temporální vlastnosti jednotlivých typů publicistických mluvených projevů. Podle jeho výzkumu televizních a rozhlasových pořadů z let 1934–1996 byly nejrychlejší mluvním tempem ze všech zkoumaných typů prosloveny přímé rozhlasové sportovní reportáže, a to tempem 5,57 slabik za sekundu. Tyto pořady měly také nejvyšší variabilitu co do mluvního tempa. Čtené zprávy dosahovaly průměrného tempa 5,08 sl./sec v případě redaktorů-mužů, jejich ženské protějšky měly při čtení zpráv průměrné tempo o něco pomalejší – 4,82 sl./sec; zde je tedy možné spekulovat i o vlivu pohlaví na mluvní tempo. Jiný typ čteného projevu – rozhlasová četba beletrie – měl oproti čteným zprávám výrazně nižší tempo. Řečové projevy v tomto typu pořadu byly ze všech zkoumaných stylů nejpomalejší s průměrným mluvním tempem 3,9 sl./sec.

Přestože další formát pořadu – beseda – byl co do typu projevu řazen mezi spontánní, mluvní tempo zde bylo pomalejší než u čtených zpráv: cca 4,3 sl./sec. Zde ale mohl hrát roli i faktor mluvčího: besed se účastnili spíše politici, vědci a ekonomové než profesionální mluvčí.

K rozdílům v tempu řeči, vyvolaným i rozdílnými mluvními styly, došla také Balkó (2005). Ve čtených projevech v jejím experimentu bylo celkové průměrné artikulační tempo 5,79 sl./sec (mluvní tempo: 4,87 sl./sec), zatímco v sice spontánních, ale kognitivně náročnějších nepřipravených monolozích bylo průměrné artikulační tempo 5,24 sl./sec (mluvní tempo: 4,26 sl./sec). V popisu schematického obrázku bytu

dosahovali mluvčí průměrného artikulačního tempa 4,9 sl./sec (mluvní tempo: 3,84 sl./sec), v popisu portréту ženy 4,94 sl./sec (mluvní tempo: 3,68 sl./sec) a v řízeném dialogu 5,76 sl./sec. K vlivu typu úlohy na artikulační tempo Balkó dodává:

*„Tempo artikulace v monologických úlohách je poznamenáno celkovým zpomalováním řeči při váhání nad formulací myšlenky, jímž mluvčí maskují potíže s jazykovým plánováním, časté je protahování koncových slabik a produkce redundantních zvuků.“* (Balkó, 2005: 193).

Druh řečové úlohy a její náročnost na stylizaci má podle autorky vliv také na celkové mluvní tempo projevu. Rovněž podíl hezitačních pauz se v jednotlivých úlohách lišil. Procentuální podíl pauz ve čtecích úlohách byl 15,3 %, v nepřipravených monolozích 20,4 %, v popisu bytu 23,1 % a v popisu portétu 22,6 %. V popisech obou obrázků – ať už schematického náčrtu bytu, nebo v propracovaném portrétu ženy – mluvčí dosahovali podobných hodnot; tato podobnost byla statisticky významná. Vliv mluvního stylu na temporální charakteristiky projevu je zde tedy zřetelně patrný.

Pro přehledné srovnání výsledků mluvního tempa v souvislosti s typem projevu a s ním souvisejícím mluvním stylem viz Bartošová, 2016: 34–35.

### **2.5.1.8 Fonotaktika češtiny**

#### **2.5.1.8.1 Stavba slabiky**

Jedním z univerzálních principů, společných pro všechny lidské jazyky, je tvoření základních jednotek souvislé řeči – slabik. Existence slabiky je podmíněna přítomností sonorní složky, tzv. *jádra* (nukleu) slabiky. V některých případech ale dochází k „relativizaci“ sonority – příkladem mohou být japonské desonorizované vokály či česká onomatopoická slova bez jediné znělé hlásky. Dojem slabičnosti je zde zachován díky temporálním proporcím jednotek; v případě slova složeného z neznělých konsonantů roli jádra slabiky přebírá nejvýraznější element (např. ve slově *pst* považujeme za jádro slabiky hlásku [s]).

Kromě jádra může mít slabika jeden či více konsonantických, nebo spíše méně sonorních elementů: *préturu* – jednu či více hlásek před jádrem – a *kodu*, jednu nebo více hlásek za ním. Nukleus a koda jsou obvykle nazývány slabičným základem, neboť se mezi těmito dvěma stavebními kameny projevují kompenzační strategie (např. v angličtině) na rozdíl od *prétury* a *nukleu*.

Elementární formou slabiky je kombinace konsonantu a vokálu v tomto pořadí. Slabika CV je charakterizována průběhem od striktury k apertuře a bývá základem jazykových inventářů.

Jádrem slabiky bývají v češtině i v řadě dalších jazyků zejména vokály, řidčeji sonorní souhlásky jako [r, l, m, n]; ve vzácných případech i šumové souhlásky (např. již zmíněná onomatopoeia typu *pst*). Préturu v češtině tvoří až pět konsonantů (*vzkvět* – [fʃkvjet]), zatímco v kodě se mohou vyskytovat maximálně tři souhlásky (*zábst* – [za:pst]). I přes tuto komplexitu je ale v češtině obecnou tendencí inklinace k tvoření slabik typu CV. Pokud se konsonant vyskytuje mezi dvěma vokály, značí vždy slabičný předěl.

Bičan (2011) uvádí jako jednu z alternativ popisů kombinatoriky hlásek v rámci české slabiky deskripci inspirovanou Kučerou (1961). Jednotlivé slabičné pozice jsou označeny jako P, N a K podle názvů pozic prétura, nukleus a koda. Číslo v dolním indexu jednotlivých pozic uvádějí pořadí fonému v prétuře či kodě, počítáno od nukleu, a počet fonémů na dané pozici. Např. ve slově *mstít* /m/ bude mít značku P<sub>33</sub>, /s/ P<sub>32</sub>, /c/ P<sub>31</sub> a /t/ K<sub>11</sub>.

Prétura					Nukleus	Koda		
P <sub>55</sub>	P <sub>54</sub>	P <sub>53</sub>	P <sub>52</sub>	P <sub>51</sub>	N	K <sub>31</sub>	K <sub>32</sub>	K <sub>33</sub>
	P <sub>44</sub>	P <sub>43</sub>	P <sub>42</sub>	P <sub>41</sub>		K <sub>21</sub>	K <sub>22</sub>	
		P <sub>33</sub>	P <sub>32</sub>	P <sub>31</sub>		K <sub>11</sub>		
			P <sub>22</sub>	P <sub>21</sub>				
				P <sub>11</sub>				

Tabulka 4. Rozvržení prétur a kod v češtině. Úprava původní verze z Bičan (2011: 271).

V práci *Phonotactics of Czech* (Bičan, 2011) dále autor vypisuje všechny fonémy, které se v pozicích naznačených v tabulce 4 mohou vyskytnout. Tento seznam je pro větší srozumitelnost upraven a v některých případech doplněn (např. jako možná jádra slabik byly přidány /m/ a /n/). Ponecháno zůstává autorovo značení tzv. *znělostních archifonémů*. Pod tímto termínem Bičan rozumí jednotky korespondující s několika fonémy a definované jako souhrn distinktivních rysů, které jsou těmto fonémům společné; znělost těchto fonémů je však dána kontextem, ve kterém se vyskytují (Bičan, 2011: 26–27). Znělostní archifonémy jsou v Bičanově úzu značeny velkým písmenem; např. archifoném /T/ zahrnuje fonémy /t/ a /d/ v závislosti na kontextu. Naopak byl

z přehledu vyřazen archifoném místa tvoření /M/, neboť je vždy realizován jako [m] (Bičan, 2011: 31) a pro účely této práce není stěžejní.

pozice	fonémy
N	/e/, /i/, /a/, /o/, /u/, /e:/, /i:/, /a:/, /o:/, /u:/, /eu/, /au/, /ou/, /r/, /l/, /m/, /n/
P <sub>11</sub>	/p/, /b/, /t/, /d/, /t̪/, /d̪/, /k/, /g/, /f/, /v/, /s/, /z/, /š/, /ž/, /x/, /h/, /m/, /n/, /ň/, /j/, /ř/, /r/, /l/
P <sub>21</sub>	= P <sub>11</sub>
P <sub>22</sub>	/p/, /b/, /t/, /d/, /k/, /g/, /f/, /v/, /s/, /z/, /š/, /ž/, /x/, /h/, /j/, /ř/, /r/, /l/, /m/, /P/, /T/, /K/, /F/, /S/, /Š/, /X/
P <sub>31</sub>	/p/, /b/, /t/, /d/, /t̪/, /d̪/, /k/, /v/, /s/, /š/, /ž/, /x/, /h/, /m/, /n/, /ň/, /j/, /ř/, /r/, /l/
P <sub>32</sub>	/p/, /b/, /t/, /d/, /k/, /f/, /v/, /s/, /z/, /š/, /ž/, /x/, /h/, /ř/, /m/, /T/, /S/, /Š/
P <sub>33</sub>	/b/, /t/, /d/, /k/, /v/, /s/, /z/, /x/, /h/, /j/, /ř/, /r/, /l/, /m/, /P/, /T/, /K/, /F/, /S/, /Š/, /X/
P <sub>41</sub>	/v/, /m/, /n/, /ň/, /j/, /ř/, /r/, /l/
P <sub>42</sub>	/p/, /b/, /t/, /d/, /k/, /v/, /s/, /š/, /x/, /h/, /m/
P <sub>43</sub>	/t/, /k/, /s/, /z/, /x/, /h/, /ř/, /T/, /S/, /Š/
P <sub>44</sub>	/t/, /h/, /l/, /m/, /P/, /T/, /F/, /S/, /Š/
P <sub>51</sub>	/j/
P <sub>52</sub>	/v/
P <sub>53</sub>	/k/
P <sub>54</sub>	/T/, /S/
P <sub>55</sub>	/F/, /S/
K <sub>11</sub>	/P/, /T/, /T̪/, /K/, /F/, /S/, /Š/, /X/, /m/, /n/, /ň/, /j/, /ř/, /r/, /l/
K <sub>21</sub>	/P/, /T/, /K/, /F/, /S/, /Š/, /X/, /m/, /n/, /ň/, /j/, /ř/, /r/, /l/
K <sub>22</sub>	/P/, /T/, /T̪/, /K/, /F/, /S/, /Š/, /X/, /m/, /n/, /ň/, /ř/, /r/, /l/
K <sub>31</sub>	/P/, /T/, /K/, /m/, /n/, /j/, /r/, /l/
K <sub>32</sub>	/P/, /T/, /K/, /S/, /Š/
K <sub>33</sub>	/T/, /K/, /S/, /Š/

Tabulka 5. Výčet fonémů v jednotlivých pozicích ve struktuře slabiky: N = nukleus, P = prétura, K = koda. Úprava původní verze z Bičan (2011: 272). Ponechán autorův původní zápis.

Ačkoli mluvčím daného jazyka obvykle nečiní problém určit počet slabik ve slově nebo ve vyšší prozodické jednotce, stanovení hranic jednotlivých slabik nemusí být jednoznačné – při přirozeném slabikování umí posluchač určit hranici slabiky pouze v případě, že se konsonanty a vokály pravidelně střídají (Ludvíková, 1987). Pro situace, kdy je mezi jednotlivými nukleji více než jeden konsonant, navrhli Kučera a Monroe (1968) několik pravidel pro zjišťování slabičných hranic: předěl by měl být stanoven tak, aby vzniklé prétury a kody nezvýšily počet prétur a kod na počátcích a koncích slov přítomných v korpusu, nebo by se mělo o hranici rozhodnout ve prospěch nejfrekventovanější prétury či kody (Kučera & Monroe, 1968: 48; s podobnými faktory pracuje i Palková, 1994: 270–271 či Bičan, 2015). V jejich práci se však nalézají i sporná

místa a specifika (autoři např. považovali neslabičné předložky a předponu vz- za samostatné jednotky), která mohou způsobit rozdíly ve srovnání s novějšími pracemi.

Podle Palkové (1994) pro určení mezislabičné hranice hraje roli zvukový kontrast sonority v řetězu hlásek, struktura frekventovaných slabik a případně i morfologický šev. Ačkoli slabika na rovině zvukové a morfém na rovině jazykové si příliš neodpovídají, morfematická stavba slova může realizaci slabiky ovlivňovat. Zřetelně je toto působení vidět na slovech obsahujících prefix před samohláskou, která může být vyslovena s hlasivkovou okluzivou, tzv. *rázem*, anebo bez ní. Přítomnost, či nepřítomnost rázu má přímý vliv na rozdělení slova na slabiky: slovo *rozoraný* může být vysloveno jako [ros.ʔo.ra.ni:], nebo jako [ro.zo.ra.ni:].

Palková (1994) uvádí také několik poznatků ohledně slabičné struktury v češtině – vždy jde ale pouze o tendence, nikoli o obecně platná pravidla: V pozici P<sub>x1</sub> se často nachází sonora či /r/ (*drát, přes*); v pozicích P<sub>22</sub> a P<sub>21</sub> se obvykle spolu nachází obstruent a sonora (*smích*); v pozicích P<sub>33</sub>, P<sub>32</sub> a P<sub>31</sub> je častou kombinací frikativa, exploziva a sonora (*strýc*); v pozici P<sub>33</sub> bývá často foném /s/ (*sklad*), kombinace dvou sonor obvykle znamená /m/ a likvidu (*mléko, mrak*).

#### 2.5.1.8.2 Konsonanticko-vokální stavba slabiky

Ludvíková (1987) zpracovávala dva české mluvené projevy (odbornou přednášku a diskusi) čítající celkem cca 10 000 slabik, z nichž vyděluje 13 konsonanticko-vokálních slabičných struktur: CV (60 %), CVC (17 %), CCV (10 %), V (5 %), CCVC (4 %), VC (2 %); typy CVCC, CCCV, CCCC, CCCVC, CCVCC, CVCCC a CCCVCC pokrývají zbývající 2 % slabičných struktur. S těmito typy se můžeme setkat v různých částech slova. Začátek slova je co do struktur nejbohatší; v materiálu Ludvíkové (1987) se objevily typy CV, CCV, V, VC, CVC, CCCV, CCCC, CCCVCC; uprostřed slova se vyskytovaly typy CV, CVC, CCV, CCVC, V, CCCVC, VC. Na konci slova Ludvíková (1987) uvádí jen šest možných typů slabik – CV, CVC, CCV, CCVC, CVCC a CCCV, přičemž nejběžnějšími slabikami na koncích slov jsou typy CV a CVC (Ludvíková, 1987: 108). Slabičné typy CV, CVC a V označují za nejčastější struktury v jednoslabičných slovech ve zkoumaném materiálu složeném celkem z 10 280 slov také Kučera a Monroe (1968).

Aleš Bičan (2015) ve svém výzkumu na Fonologickém lexikálním korpusu češtiny došel co do frekvence jednotlivých slabičných typů k podobným výsledkům jako Ludvíková – v analýze domácí slovní zásoby, reprezentované téměř 46 000 fonologickými slovy a bezmála 147 000 slabikami, označil za nejfrekventovanější český slabičný typ CV (48,05 %), následují typy CVC (18,17 %), CCV (17,11 %), CCVC (5,51 %) a CVCC (3,13 %). Je třeba však připomenout, že Bičan (2015) ve své analýze považoval české afrikáty za kombinaci dvou konsonantů, což mohlo jeho výsledky částečně ovlivnit.

Z výše uvedených skutečností vyplývá, že ačkoli je čeština známa i svými komplexními souhláskovými shluky zejména v préture slabiky, většina slabik ve standardním projevu se obejde bez konsonantických skupin předcházejících jádro slabiky. S tím souhlasí i další výsledky Kučery a Monroa (1968), které shrnují nejfrekventovanější slabičné préture a kody. Nejčastější préturem je typ C (63 % z préture v korpusu), následuje CC (24 %) a nulová préture (12 %). Komplexnější shluky konsonantů v préture jsou vzácné – nejčastějšími konsonantickými skupinami jsou shluky dvojčlenné (např. Mazlová, 1946; Volín & Churaňová, 2010). Kody jsou podle Kučery a Monroa (1968) nejčastěji nulové (73 %), popř. je tvoří jeden konsonant (26 %); shluky CC a CCC v kodě se v materiálu autorů studie vyskytují zřídka.

Churaňová (2013) provedla podobnou analýzu jako Kučera a Monroe na jednoslabičných mluvních taktech ve zpravodajských textech. Na rozdíl od výše uvedených autorů došla k závěru, že nejčastějším typem jednoslabičných taktů je relativně komplexní vzorec CCVC – rozdíl však lze vysvětlit tím, že neanalyzovala slova, ale právě mluvní takty, tedy jednoslabičná slova stojící samostatně a nesoucí přízvuk. Zatímco mezi jednoslabičnými slovy se v českých textech vyskytuje řada předložek a spojek typu CV a CVC, nestávají se obvykle tyto slovní druhy samostatnými takty. Těmi byla v daném materiálu spíše plnovýznamová slova, jako např. *dnes*.

Nejfrekventovanějšími slabikami v mluvených projevech zkoumaných Ludvíkovou byly slabiky /di, ní, to, a, ta, po, je, ně, ja, na, vá, ta, ne, ka, no, se, pro, tu, ko, o/ (Ludvíková, 1987: 106; zápis ponechán dle původní studie). Autorka uvádí, že na frekvenci těchto slov má vliv skutečnost, ve kterém slově a slovním druhu se nacházejí, a rovněž frekvence těchto slov v projevu. Na četnost výskytu daného slova pak může mít vliv i např. téma projevu, individuální mluvní návyky (např. časté užívání výplňkových slov)



a také stupeň formálnosti řeči. Nejfrekventovanější slabiky podle pozice ve slově jsou následující: /te, po, ne, pro, a, ob, ta, ja, vi, u/ na začátku slova, /vá, no, ja, jme, ka, jek, ko, zo, li, zí/ jsou slabiky typické pro mediální pozici a /ňí, ci, ho, tu, te, ňe, ká, ré, né, ké/ se vyskytovaly na koncích slov (Ludvíková, 1987: 107).

#### 2.5.1.8.3 Frekvence hlásek a jejich kombinací v češtině

Podle Ludvíkové (1987) jsou nejpoužívanějšími fonémy či hláskami v češtině krátké vokály (32,12 %), neznělé obstruenty (22,95 %) a sonory spolu s /r/ (22,35 %), méně frekventované jsou znělé obstruenty (13,4 %) a dlouhé vokály (8,29 %). Nejnižší frekvenci užívání má diftong /ou/ (0,89 %).

Poměr samohlásek oproti souhláskám v českých souvislých spisovných textech se jeví napříč jednotlivými studiemi jako pozoruhodně stabilní – cca 59 % hlásek v textech z různých stylů tvoří konsonanty a zhruba 41 % vokály.

	<b>Mazlová (1946)</b>	<b>Kučera (1963)</b>	<b>Ludvíková &amp; Kraus (1966)</b>	<b>Bartoň et al. (2009) hlásky – tokeny</b>	<b>Bartoň et al. (2009) hlásky – typy</b>	<b>Churaňová (2012)</b>
C	59,30 %	58,48 %	58,70 %	58,54 %	58,25 %	58,25 %
V	40,70 %	41,52 %	41,30 %	41,46 %	41,75 %	41,75 %

Tabulka 6. Srovnání poměru konsonantů a vokálů v českých souvislých textech. Kromě Churaňové (2012), která zkoumala výhradně publicistické útvary, všechny studie obsahovaly texty různých stylů.

Vzhledem ke stylu projevu tento poměr může proměňovat – Mazlová (1946) například zjistila, že v poezii z jejího materiálu se překvapivě vyskytuje méně samohlásek (39,67 %) než v próze (41,47 %). K podobnému závěru došli i Ludvíková a Kraus (1966), kteří zjistili nejvyšší poměr vokálů v odborných textech (42 %), zatímco konkrétně v Hrubínově poezii bylo samohlásek nejméně (39 %).

Kvantitativními vlastnostmi českých fonémů a hlásek se systematicky zabývala Mazlová (1946), dále Kučera (1963), Ludvíková a Kraus (1966) a nejnovější přehled o frekvenci českých hlásek přinesla publikace *Statistiky češtiny* (Bartoň et al., 2009). Všechna zmíněná zkoumání frekvence českých hlásek v souvislých textech přinesla srovnatelné výsledky. Tabulka 7 uvádí statistiky jednotlivých hlásek či fonémů v českých souvislých textech sestupně podle frekvence. Ačkoli Mazlová (1946) a Bartoň et al. (2009) používají fonetickou transkripci a pracují s hláskami a Kučera (1963) a Ludvíková a Kraus (1966) operují s transkripcí fonologickou a s fonémy, výsledky jsou relativně podobné.

Nejfrekventovanějším fonémem či hláskou v češtině je /ε/ – v této skutečnosti se všechna uvedená měření shodnou. Následují vokály /o, a, ɪ/ v tomto pořadí u Kučery (1963), Ludvíkové a Krause (1966) a Bartoň *et al.* (2009); v materiálu Mazlové (1946) bylo /a/ frekventovanější než /o/. Dalšími častými fonémy/hláskami jsou /s, l, n, t/ – v každém z výzkumů v pozmeněném pořadí, vždy však mezi 4–6 % výskytů ze všech hlásek. Relativně frekventované jsou také fonémy /i:, v, k, m, r/. Frekvence dalších fonémů a hlásek se liší v závislosti na dané studii; mezi méně frekventované ale vesměs patří diftongy a fonémy /o:, u:, g, ʃ/.

Mazlová (1946)		Kučera (1963)		Ludvíková & Kraus (1966)		Bartoň <i>et al.</i> (2009)		Bartoň <i>et al.</i> (2009)	
hláska	%	foném	%	foném	%	hláska – tokeny	%	hláska – typy	%
ε	9,7	ε	9,4	ε	9,79	ε	9,43	ε	8,26
a	7	o	8,24	o	6,91	o	7,24	o	7,97
o	6,5	a	6,83	a	6,66	a	6,8	a	7,16
ɪ	6,1	ɪ	6,49	ɪ	6	ɪ	6,26	ɪ	7,1
s	5,2	l	5,58	t	4,76	s	4,85	n	5,17
l	5,1	s	4,65	s	4,71	t	4,83	r	4,87
n	4,4	t	4,31	n	4,56	l	4,54	l	4,63
t	4,4	n	4,17	l	4,35	n	4,3	s	4,35
i:	4,1	i:	3,7	k	4,02	i:	3,72	v	4,22
v	3,9	v	3,7	i:	4	v	3,67	t	4,14
k	3,7	k	3,66	v	3,89	k	3,55	k	4
m	3,6	m	3,66	m	3,71	r	3,51	m	3,72
r	3,2	r	3,12	r	3,66	m	3,37	i:	3,23
j	3,1	u	3,07	p	3,11	p	3,3	p	2,86
p	2,9	j	2,98	j	2,89	d	2,68	u	2,77
d	2,8	p	2,98	u	2,76	ʔ	2,57	d	2,07
a:	2,3	d	2,97	d	2,74	j	2,37	a:	2,01
u	2,3	ɲ	2,18	a:	2,4	u	2,37	ts	1,76
ɲ	1,9	a:	2,08	ɲ	1,99	ɲ	2,13	j	1,63
z	1,9	z	1,74	z	1,94	a:	2,09	z	1,52
b	1,7	h	1,48	b	1,83	z	1,74	b	1,51
h	1,6	ʃ	1,44	ts	1,35	b	1,57	h	1,5
ʃ	1,6	b	1,35	h	1,27	ts	1,44	ɲ	1,46
x	1,4	ts	1,3	ε:	1,25	h	1,29	ʔ	1,44
ts	1,3	tʃ	1,22	ʃ	1,24	ʃ	1,15	x	1,24
ε:	1,2	ɹ	1,16	x	1,2	ε	1,08	ε:	1,12
c	1,2	x	1,15	ɹ	1,19	x	1,03	f	0,99
ɹ	1,1	ε:	1,11	tʃ	0,96	c	0,94	c	0,96
tʃ	0,9	c	1,11	c	0,87	tʃ	0,91	ʃ	0,93
ʒ	0,9	ʒ	0,91	ou	0,87	f	0,79	g	0,85

ou	0,9	f	0,61	z	0,85	z	0,79	tj	0,85
f	0,7	u:	0,6	f	0,73	ř	0,72	ou	0,77
u:	0,6	ť	0,59	u:	0,59	ou	0,69	ř	0,53
ť	0,5	g	0,46	ť	0,48	u:	0,61	u:	0,49
g	0,3	o:	—	g	0,4	ť	0,51	ť	0,37
o:	—	au	—	o:	0,05	ř	0,48	z	0,36
au	—	eu	—	au	0,02	g	0,44	ř	0,34
eu	—	ou	—	eu	—	η	0,12	η	0,32

Tabulka 7. Zastoupení fonémů a hlásek v materiálech Kučery (1963), Mazlové (1946), Ludvíkové a Krause (1966) a Bartoně *et al.* (2009). Sloupec *hláska – tokeny* reprezentuje četnost výskytu každé hlásky v celém korpusu SYN2005; sloupec *hláska – typy* reprezentuje počet hlásek v jedinečných slovních tvarech v témže korpusu. Foném /r/ se u Mazlové (1946) vyskytoval v 0,6 % ve znělé a v 0,5 % v neznělé variantě. Hodnoty Mazlové a Kučery značí průměrný výskyt hlásek na 1000.

au	0,05	w	0,21
o:	0,03	au	0,13
w	0,02	o:	0,09
dz	0,01	dz	0,03
dž	0,01	dž	0,03
/ü/	0	/ö/	0,02
/ö/	0	/ü/	0,02
γ	0	/ä/	0,02
/ä/	0	γ	0,01

Ludvíková a Kraus (1966) se dále zabývají kombinačními možnostmi jednotlivých fonémů. Hlávky či fonémy, které mohou segmentu předcházet, nebo ho následovat, tvoří tzv. *valenční pole fonému*. Hodnota předního valenčního pole udává, kolik fonémů může daný foném předcházet, a hodnota zadního valenčního pole naopak vypovídá o tom, kolik fonémů může za daným segmentem následovat.

22 b 22	18 h 18	32 j 12	27 t 28
25 ts 18	18 x 26	21 p 27	24 c 11
23 tj 17	18 j 35	29 r 34	28 v 25
20 d 25	28 k 24	22 ř 20	21 z 25
17 ě 7	33 l 33	25 s 29	21 ž 18
16 f 19	28 m 29	22 ř 21	36 # 36
11 g 17	32 n 31		

Tabulka 8. Přední a zadní valenční pole českých konsonantických fonémů (Ludvíková & Kraus, 1966). Symbol # představuje slovní hranici.

31 a 31	29 i 29	29 u 28
28 a: 25	26 i: 26	22 u: 23
29 ε 35	31 o 33	19 ou 22
18 ε: 23	13 o: 9	4 au 5

Tabulka 9. Přední a zadní valenční pole českých vokálních fonémů (Ludvíková & Kraus, 1966).

Valenční pole dosahuje nejvyšších možných hodnot u slovní hranice (36 # 36). Šíří kombinačních možností se oběma těmito hodnotám zároveň nejvíce blíží sonora /l/ (33 l

33), samohlásky /o/ (31 o 33) a /a/ (31 a 31) a nazála /n/ (32 n 31). Vysokou hodnotu předního valenčního pole má také foném /ɲ/ (32); hodnota zadního valenčního pole je nejvyšší u vokálu /ɛ/ (35) a u sonor /j/ (35) a /r/ (34). Palatály mají nejvyšší rozdíly mezi hodnotami předního a zadního valenčního pole, a to vždy ve prospěch předního. Nejméně se v češtině kombinují vokály /âu/ a /o:/ a konsonanty /ʃ/ a /g/; omezenější možnosti kombinace mají také obstruenty /c/, /f/ a /h/.

Jednou ze skutečností týkající se fonotaktické struktury češtiny je fakt, že česká slova typicky začínají konsonanty a končí vokály, což poukazuje také na snahu o produkci otevřených slabik (CV), jak ověřila např. Churaňová (2012). Studie Ludvíkové a Krause (1966) a Bartoně *et al.* (2009) dospěly k velmi podobným závěrům, co se týče nejfrekventovanějších fonémů či hlásek na začátcích a koncích slov.

	začátek slova			konec slova		
	Ludvíková & Kraus, 1966	Bartoň <i>et al.</i> , 2009: tokeny	Bartoň <i>et al.</i> , 2009: typy	Ludvíková & Kraus, 1966	Bartoň <i>et al.</i> , 2009: tokeny	Bartoň <i>et al.</i> , 2009: typy
1.	s	s	p	ɛ	ɛ	ɪ
2.	p	p	n	ɪ	ɪ	m
3.	n	n	s	a	a	a
4.	v	v	k	i:	i:	ɛ
5.	j	a	v	o	o	i:
6.	t	j	m	m	m	o
7.	z	t	b	u	t	u
8.	m	m	z	ɛ:	u	x
9.	a	z	r	t	ɪ	ɛ:
10.	k	k	o	k	ɛ:	a:

Tabulka 10. Nejfrekventovanější fonémy/hlásky na začátku a konci slova, srovnání studií Ludvíkové a Krause (1966) a Bartoně *et al.* (2009). Sloupec *tokeny* reprezentuje četnost výskytu každé hlásky v celém korpusu SYN2005; sloupec *typy* reprezentuje počet hlásek v jedinečných slovních tvarech v témže korpusu.

Bartoň *et al.* (2009) kromě nejčastějších hlásek podle jejich pozice ve slově v celém korpusu také zpracoval frekvenci hlásek v lexikonu daného materiálu – každý slovní tvar, který se v něm vyskytl, započítal pouze jednou. Tato analýza měla jiné výsledky než celková statistika (např. nejfrekventovanější hláskou na začátku slova nebylo [s] jako v celkové analýze, nýbrž [p], a na konci slova se ukázalo jako nejčastější [ɪ] místo celkově nejpoužívanějšího [ɛ]), což dokazuje vliv frekventovaných slov a slovních tvarů na celkovou fonotaktickou stavbu projevu. První desítka nejčastějších hlásek na začátku a konci slova je uvedena v tabulce 10.

Nejčastějšími dvojčlennými skupinami hlásek v češtině jsou spojení konsonantu a vokálu CV (např. /je, ne, na, po, ní/; zápis dle Bartoně *et al.*, 2009). Nejobvyklejší dvojicí konsonantů CC je /st/, která se vyskytuje v řadě českých slov jako prétura slabiky (např. *stavět, starý, stát, stékat...*). Kombinace vokálu a konsonantu VC (např. /ov/) není tak obvyklá jako dvojice s opačným pořadím hlásek, ale objevuje se v českých projevech častěji než kombinace dvou konsonantů. V těchto faktech se studie Ludvíkové a Krause (1966), Bartoně *et al.* (2009) a Churaňové (2012) shodnou. Typ VV (např. /eu/) je dle Churaňové (2012) a Bartoně *et al.* (2009) méně obvyklý a v korpusu SYN2005 se objevuje jen u substantiv, adjektiv a sloves. Churaňová (2012) rozlišovala v rámci této analýzy i slabikotvorné likvidy, a tak dalšími dvojicemi, které se v jejím materiálu častěji objevovaly v libovolné pozici, byly např. dvojice CR a RC. Všechny uvedené kombinace konsonantů a vokálů byly rovněž zaznamenány v analýze taktů (Churaňová, 2012). Kromě uvedených kombinací byla v materiálu Churaňové (2012) zaznamenána dvojice RV v rámci taktu – toto spojení vzniklo díky výslovnosti za sebou jdoucích slov *sedm až* a jejich sloučením do jednoho mluvního taktu.

Trojčlenné skupiny byly v korpusu SYN2005 i v materiálu Churaňové (2012) typu CVC (např. /ter, val, bil, nos, jak/; zápis dle Bartoně *et al.*, 2009), VCV (např. /ova, eňí, ého, ale, ali/), CCV (např. /vje, pro, mňe, sta, pře/) a VCC (např. /ost, ick, est, ist, ost'/). Trojice typu CCC (např. /str/), CVV (např. /neo/) a VVC (např. /eut/) byly v korpusu Bartoně *et al.* (2009) vzácnější; Churaňová (2012) připojuje k těmto trojicím, které mají relevantní výskyt u slov i taktů, také CRC se slabikotvornou likvidou a dále velmi zřídka se vyskytující trojice CCR, RCC, VCR a RCV. Autoři *Statistik češtiny* i Churaňová (2012) rovněž uvádějí trojici vokálů VVV v rámci slov – tato kombinace mohla vzniknout výslovností zkratk nebo přítomností cizích slov.

Jak již bylo naznačeno, v iniciální pozici se setkáme nejčastěji s dvojicí hlásek typu CV (např. /po, se, ne, na, je/). Kombinace dvou konsonantů je v češtině na začátku slova o něco méně frekventovaná (např. /pr, př, st, sp, kt/); na rozdíl od materiálu Ludvíkové a Krause (1966) je v korpusu SYN2005 a v materiálu Churaňové (2012) dvojice VC (např. /ot/) na začátku slova méně obvyklá než CC. Dvojice VV se v těchto korpusech na začátku slova vyskytovala minimálně, podobně jako CR v materiálu Churaňové (2012).

	Ludvíková & Kraus, 1966	Bartoň <i>et al.</i> , 2009: tokens	Bartoň <i>et al.</i> , 2009: typy
1.	je	je	ov
2.	st	st	ne
3.	ne	po	ro
4.	na	ňí	en
5.	po	na	an
6.	se	ov	va
7.	ňí	ro	er
8.	ro	ne	vi
9.	ňe	se	st
10.	en	ňe	al
11.	le	li	li
12.	em	le	ko
13.	la	la	ra
14.	ov	al	ím
15.	li	te	sk
16.	to	to	la
17.	ko	en	te
18.	te	el	el
19.	el	pr	le
20.	pr	il	no

Tabulka 11. Nejfrekventovanější dvojice fonémů/hlásek, srovnání studií Ludvíkové a Krause (1966) a Bartoně *et al.* (2009). Sloupec *tokens* reprezentuje četnost výskytu každé hlásky v celém korpusu SYN2005; sloupec *typy* reprezentuje počet hlásek v jedinečných slovních tvarech v témže korpusu. Ve sloupci *tokens* byly původně zařazeny také dvojice /ʔa/ a /ʔo/, které v původních datech byly 3. a 10. nejfrekventovanější (v tomto pořadí), zde však ráz není považován za hlásku. Studie Ludvíkové a Krause (1966) s rázem jakožto fonémem nepracuje, a tak jsou pro lepší srovnání těchto studií obě dvojice s rázem vynechány. Zápis byl ponechán dle původních studií.

Nejčastější trojicí v iniciální pozici byla v korpusu SYN2005 a rovněž v korpusu Churaňové (2012) skupina hlásek typu CVC (např. /jak/, /bil/, /pod/, /neb/, /jed/). Další frekventovanou skupinou se ukázala být CCV (např. /pro/, /pře/, /kte/, /při/, /mňe/). Méně zastoupené byly typy VCC (např. /alk/), VCV (např. /oko/), CVV (např. /neu/) a VVC (např. /eut/), v materiálu Churaňové (2012) také trojice CCC, CRC, CCR a VCR (pro slova i takty). Velmi vzácnou trojicí byl v korpusu SYN2005 pak typ VVV.

Ve finální pozici byla v korpusu SYN2005 i v materiálu Churaňové (2012) nejčastější dvojice CV (např. /ňí, se, na, je, ho/), relativně frekventovaná byla i kombinace VC (např. /em, ích, ím, il, ak/). Méně zastoupeny byly skupiny CC (např. /st/) a VV (např. /ie/; na epentetické /j/ se v analýze Bartoně *et al.* (2009) zřejmě nebral zřetel). Další méně zastoupené skupiny v korpusu Churaňové (2012) byly CR a RC (pro slova i takty).

Na koncích slov se v SYN2005 i v analýze slov a taktů Churaňové (2012) obvykle vyskytovala trojice hlásek typu VCV (např. /eňí, ého, ila, ale, áňí/), silně zastoupeny byly i skupiny CVC (např. /ňím, vat, ních, ňích, val/), CCV (např. /st'i, vje, ské, vňí, tňí/). Méně zastoupena byla skupina VCC (např. /ost/) a velmi vzácné byly typy CVV (např. /die/ –

ve *Statistikách češtiny* opět nebylo bráno v potaz epentetické /j/), VVC (např. /uet/), CCC (např. /mst/) a VVV (pouze v SYN2005). Dalšími méně frekventovanými kombinacemi byly v materiálu Churaňové (2012) trojice CRC, CCR, VCR a RCV; v případě taktů také skupina RVC (výše zmíněný takt složený ze slov *sedm až*).

V materiálu Churaňové (2012) se také objevilo několik desítek případů kombinace CCCC a jedna skupina CCCCC v rámci analýzy slov i taktů (výraz *čerstvě*).

#### 2.5.1.8.4 Konsonanticko-vokální stavba slov a mluvních taktů

Výzkum Churaňové (2012, 2013) se zabýval konsonanticko-vokálními vzorci základních jednotek na rovinách jazyka a rytmu. Materiál, sestávající ze čtených zpravodajských relací, obsahoval 6639 slov a 5368 mluvních taktů; různých CV vzorců se v rámci slov objevilo 413 a v rámci taktů 559, přičemž ráz byl považován za samostatnou jednotku nefonematického charakteru. Mluvních taktů bylo méně než slov, neboť v rámci jednoho taktu se může objevit několik slov, zatímco opačná situace, kdy slovo bylo rozděleno dvěma realizovanými přízvuky na dva mluvní takty, byla velmi vzácná. Této skutečnosti také odpovídá větší množství typů taktových vzorců.

SLOVA			TAKTY		
vzorec	výskytů	%	vzorec	výskytů	%
CV	603	9,08	CVCV	380	7,08
CVCV	481	7,25	CVCCV	296	5,51
CVCCV	354	5,33	CVCVCV	229	4,27
C	278	4,19	CCVCV	223	4,15
CVCVCV	271	4,08	CVCVC	209	3,89
CCVCV	267	4,02	CVCCVCV	169	3,15
CVCVC	265	3,99	CCVCVCV	150	2,79
ʔV	237	3,57	CVCVCVC	132	2,46
CVC	211	3,18	CCVCCV	125	2,33
CVCCVCV	178	2,68	CVCVCVCV	124	2,31

Tabulka 12. Nejfrekventovanější fonotaktické vzorce u slov a taktů a přepočet na procenta. 100 % = celkový počet slov (6639) a taktů (5368). Převzato z Churaňová, 2013.

V korpusu byly vysledovány preference pro četnost výskytu slov i taktů s určitou fonotaktickou strukturou. Nejčastějším CV vzorcem českého slova byla prostá kombinace konsonant-vokál (CV, např. [ʒɛ]), následovaly struktury CVCV (např. [sɪtʃɛ]) a CVCCV (např. [dalʃi:]). Oproti tomu nejfrekventovanější strukturou českého taktu byl

vzorec CVCV, často se objevovaly i CVCCV a CVCVCV (např. [neɦoda]). Tabulka 12 ukazuje výskyt nejčastějších struktur u slov i taktů v rámci zkoumaného materiálu.

Ve zkoumaném korpusu zpravodajských relací byla nejfrekventovanějším slovním druhem substantiva (39,3 % korpusu). Značný výskyt zaznamenala i adjektiva (13,8 %), předložky (13 %) a slovesa (12,5 %). Nejmenší podíl v textech měly částice (necelé 1 %); citoslovce se v textech neobjevila. Nejčastější konsonanticko-vokalický vzorec pro každý slovní druh a variabilita CV struktur jsou zachyceny v tabulce 13.

slovní druh	nejfrekventovanější vzorec	počet výskytů	počet unikátních vzorců	variabilita vzorců
substantiva	CVCV	2607	258	9,9
adjektiva	CVCCV	915	205	22,4
předložky	CV	861	24	2,79
slovesa	CVCV	831	116	13,96
příslovce	CVCV	381	47	12,34
zájmena	CV	348	15	4,31
spojky	ʔV	322	18	5,59
číslovky	CCV	304	52	17,11
částice	ʔVC, CVC	63	17	26,98
citoslovce	-	0	0	0

Tabulka 13. Nejčastější vzorce pro dané slovní druhy, počet vzorců a jejich variabilita. Úprava původní verze z Churaňová, 2013. Sloupec *variabilita vzorců* obsahuje poměrné vyjádření variability struktur vzhledem k počtu všech výskytů daného slovního druhu v korpusu. Tato hodnota je pro větší přehlednost vynásobena 100 a vyjadřuje počet unikátních vzorců na 100 výskytů slovního druhu, tzn. čím blíže je daná hodnota 100, tím je variabilita vyšší.

Churaňová (2013) dospěla co do nejfrekventovanějších vzorců pro konkrétní slovní druhy k velmi podobným závěrům jako Bartoň *et al.* (2009) – významnější rozdíl byl zaznamenán pouze u příslovcí, jejichž nejčastější strukturou dle *Statistik češtiny* byla CVC. Ostatní rozdíly byly pravděpodobně způsobeny jiným pojetím rázu – autoři *Statistik* ho zřejmě považovali za konsonant.

Nejvíce unikátních konsonanticko-vokalických vzorců měla v materiálu Churaňové (2012, 2013) substantiva, adjektiva a slovesa. Ačkoli substantiva a adjektiva zaznamenala podobný počet jedinečných vzorců, adjektiva vykazovala výrazně vyšší variabilitu (22,4 jedinečných CV vzorců na 100 výskytů slovního druhu) než substantiva (9,9 vzorců na 100 výskytů slovního druhu). Nejvyšší variabilita typů CV struktur byla zaznamenána u částic (26,98 vzorců na 100 výskytů slovního druhu), které však v korpusu tvořily necelé 1 % všech slovních druhů. Dalšími relativně variabilními



slovními druhy byly číslovky (17,11 vzorců na 100 výskytů); o něco méně variabilní byla slovesa (13,96 vzorců na 100 výskytů) a příslovce (12,34 vzorců na 100 výskytů). Naopak velmi málo unikátních CV vzorců vzhledem k zastoupení slovního druhu bylo nalezeno u spojek (5,59 vzorce na 100 výskytů slovního druhu) a zájmen (4,31 vzorců na 100 výskytů); zcela nejnižší variabilitu vykazovaly předložky (2,79 vzorce na 100 výskytů).

## 3 Metoda

### 3.1 Materiál

K hlavnímu experimentu bylo využito 14 čtených projevů hlasatelů Českého rozhlasu, tzn. profesionálních rodilých mluvčích češtiny. Všechny projevy byly vybrány z běžných zpravodajských relací z let 2001–2005. Texty obsahovaly celkem 7632 slov, 6216 mluvních taktů a 750 nádechových úseků. Mezi 14 mluvčími bylo 7 mužů a 7 žen; od každého mluvčího byla zpracována právě jedna nahrávka. Nahrávky pocházely z Pražského fonetického korpusu (Skarnitzl, 2010).

Z časového hlediska se projevy pohybovaly v rozmezí 1,8 až 4,6 minut; průměrné trvání jedné relace bylo 3,7 minut. Jednotlivé nahrávky obsahovaly od 234 po 603 mluvních taktů a od 268 do 766 slov; průměrný počet taktů a slov byl 444 a 545 (v tomto pořadí). Pro texty byl charakteristický zpravodajský mluvní styl, bez expresivních či dialektově příznakových prvků.

Tabulka 14 obsahuje informace o jednotlivých nahrávkách:

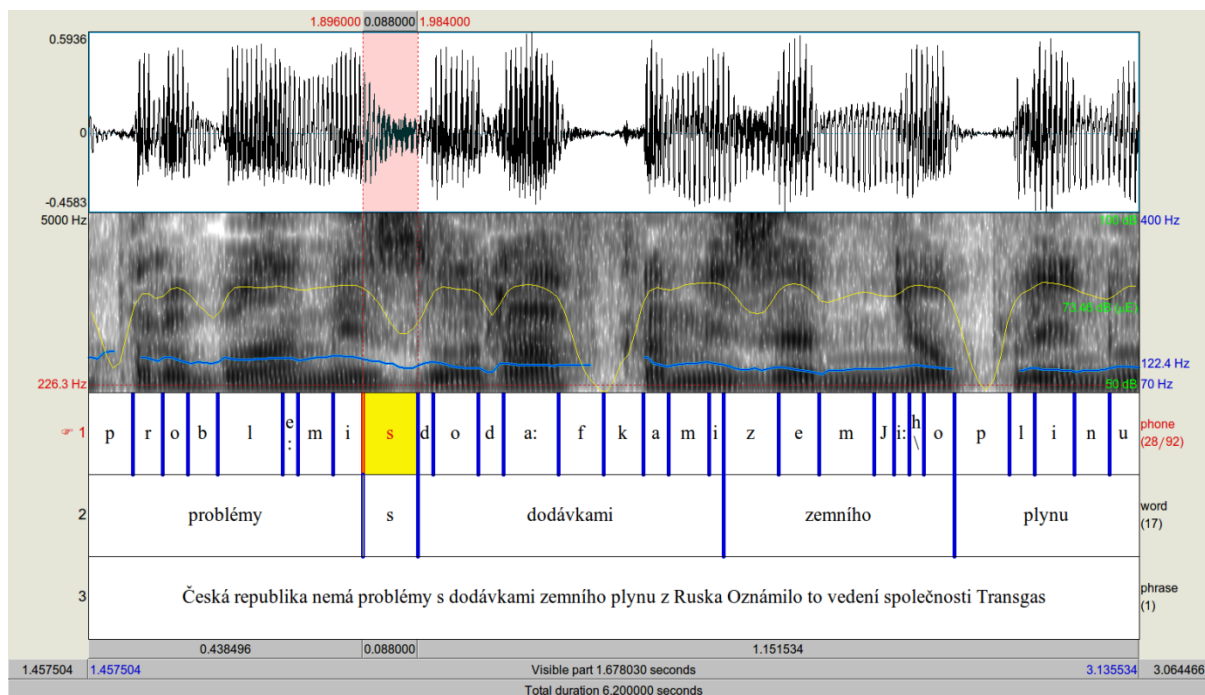
mluvčí	trvání (min.)	nádech. úseky	mluvní takt	slova
BMB01	1,8	24	234	268
CMA01	3,65	54	390	479
DYD01	4,59	69	600	720
GVA05	4,52	63	503	620
JPC04	3,84	47	450	548
KOB02	4,08	60	491	600
MSA03	3,57	45	379	460
ODC05	4,54	65	568	695
PRK01	2,64	35	327	420
SSA02	3,28	55	395	496
STR04	4,53	70	603	766
ULR05	3,75	60	416	527
VOZ01	4,17	51	519	620
WSS05	2,87	52	341	413
celkem	51,85	750	6216	7632
průměr	3,7	53,57	444	545,14

Tabulka 14. Popisné atributy nahrávek využitých v experimentu. První sloupec obsahuje kód mluvčího, barva indikuje pohlaví mluvčího (červená – žena, modrá – muž). V dalších sloupcích je uvedeno trvání nahrávek v minutách, počet nádechových úseků, mluvních taktů a slov v nahrávce.

### 3.1.1 Příprava materiálu

Část řečového materiálu byla využita již ve studii Churaňové (2012). K těmto nahrávkám byly přidány další, které byly zpracovány shodně se stávajícími (Churaňová, 2012: 79–90). Nahrávky byly nejprve rozděleny do souborů podle nádechových úseků. Ke zvukům byly poté v programu Praat (Boersma & Weenink, 2018) vytvořeny textové soubory (dále textgridy), do nichž byla posléze zanesena automatická anotace hlásek a slov pomocí souboru programů Prague Labeller (Pollák, Volín & Skarnitzl, 2007). Poté již probíhala manuální anotace na reálné mluvní takty pomocí akustických a percepčních vodítek, kterými byly např. pauza, hezitace, změny tempa nebo ukončující melodém, změna průběhu  $f_0$ , změna dynamiky, přítomnost rázu apod. (Churaňová, 2012: 81). V původních anotacích se označovaly pouze hranice mluvních taktů a neopravovaly se nepřesnosti vzniklé automatickou transkripcí (např. chybějící asimilace znělosti na hranicích slov, nezachycené rázy) a hranice hlásek, protože předmětem zmiňovaného výzkumu byly pouze konsonanticko-vokální struktury slov a mluvních taktů. Předkládaná práce se však soustředí i na konkrétní hláskovou náplň těchto fonotaktických vzorců a na trvání jednotlivých segmentů, a proto bylo nutné ve všech nahrávkách manuálně upravit hranice všech hlásek v souladu s pravidly uvedenými v Machač & Skarnitzl (2009) a opravit všechny chyby v automatické anotaci týkající se asimilace znělosti a zápisu rázů.

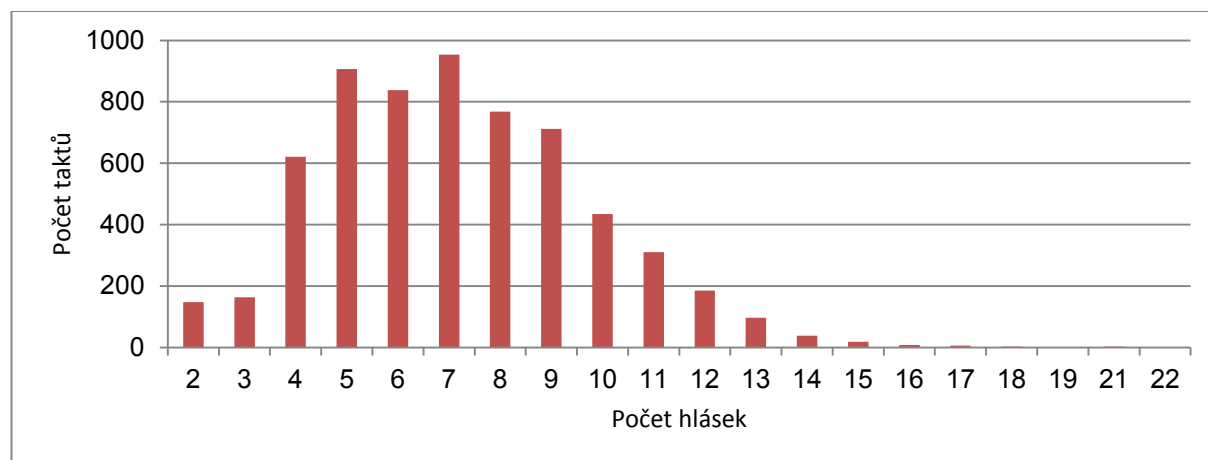
Po manuálních editacích všech textových souborů bylo možné aplikovat příkazový soubor (skript), který k hláskám přiřadil do další vrstvy v textgridu označení konsonantu, vokálu, slabikotvorné likvidy nebo rázu. Dalším skriptem byla získána všechna potřebná temporální data o taktech a hláskách: konsonanticko-vokální vzorec taktu, počet hlásek a slov v taktu, jeho hláskové složení, trvání taktu i jednotlivých hlásek v něm. Na rozdíl od výzkumu popsaného v Churaňová (2012) nebyla brána v potaz slova, a to z důvodu, že základní jednotkou, na níž se v řeči realizuje rytmus, je takt, a nikoli slovo. Na základě získaných dat byla tedy provedena základní analýza materiálu, na níž se zakládal hlavní experiment.



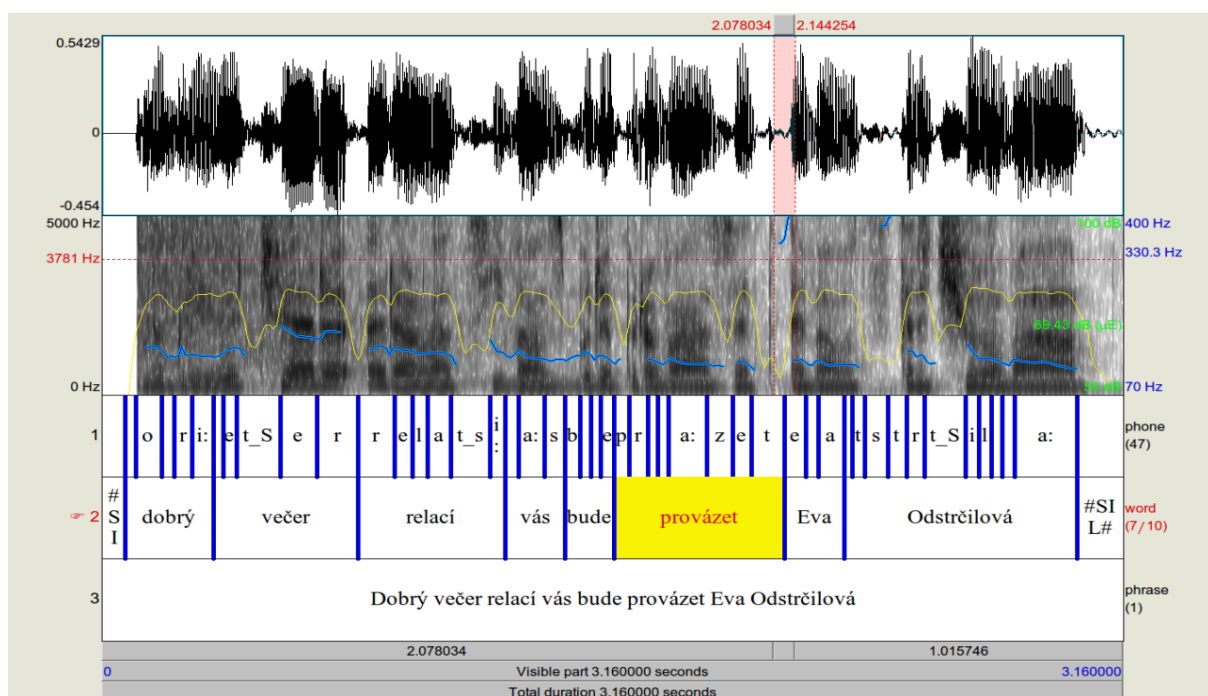
Obr. 1. Chybné automatické označení hlásky [z], vzniklé asimilací znělosti, jako neznělé [s] v souboru JPC04-02-01. Soubor rovněž obsahuje hlásku [l] s nesprávně stanovenými hranicemi ve slově *problémy*.

### 3.1.2 Fonotaktický popis materiálu

V rámci experimentu byly využity poznatky plynoucí z práce Churaňové (2012) o fonotaktických vzorcích a do experimentu zahrnuty ty vzorce, které jsou v mluveném projevu nejčastější. Celkem bylo v textech nalezeno 696 různých konsonanticko-vokalických vzorců (s rázem a slabikotvornou likvidou se počítalo jako se speciálními prvky); po odstranění vzorců s hezitacemi se jejich počet zmenšil na 635. Celkový počet hlásek včetně rázu v materiálu byl 44 493. Jak je možné vidět v grafu 2 a v tabulce 15, nejčastěji se v projevech vyskytovaly taktý složené ze sedmi hlásek (15,4 % všech taktů).



Graf 2. Četnost taktů podle počtu hlásek (se zahrnutím rázu) v nich obsažených.



Obr. 2. Vynechání anotace rázu před slovem *Eva* v souboru ODC05-01-01. V automatické anotaci rovněž nebyla zohledněna asimilace znělosti přes hranici slova u hlásky [s] ve slově *vás*.

počet hlásek	počet taktů	%
2	148	2,38
3	163	2,62
4	621	9,99
5	907	14,59
6	838	13,48
7	954	15,35
8	768	12,36
9	712	11,45
10	435	7
11	310	4,99
12	185	2,98
13	97	1,56
14	38	0,61
15	18	0,29
16	8	0,13
17	6	0,1
18	3	0,05
19	1	0,02
21	3	0,05
22	1	0,02
celkem	6216	100

Tabulka 15. Četnost taktů s určitým počtem hlásek (se zahrnutím rázu).

Nejfrekventovanějšími vzorci v materiálu byly CVCV (6,7 % ze všech taktů v materiálu), CVCCV (5,3 %) a CVCVCV (4,4 %). Následovaly vzorce CCVCV (4 %) a CVCVC (3,5 % ze všech taktů). Tyto výsledky jsou tak v souladu s nálezy studie Churaňové (2012). Prvních 20 nejčastějších konsonanticko-vokalických vzorců je zachyceno v tabulce 16.

konsonanticko-vokalický vzorec mluvního taktu	frekvence výskytu	%
<b>CVCV</b>	414	6,66
<b>CVCCV</b>	329	5,29
CVCVCV	271	4,36
<b>CCVCV</b>	250	4,02
CVCVC	217	3,49
CVCCVCV	184	2,96
CCVCVCV	163	2,62
CVCVCVCV	151	2,43
CVCVCVC	150	2,41
CVCVCCV	145	2,33
CCVCCV	139	2,24
CVCCVC	103	1,66
CCVCVC	93	1,5
CVCCVCVCV	84	1,35
CVCVCCVCV	82	1,32
CCVC	78	1,25
CVCVCVCCV	75	1,21
CVC	75	1,21
CVCCVCCV	74	1,19
ʔV	74	1,19
CCVCCVCV	71	1,14
CCVCVCVCV	69	1,11
CV	68	1,09
CCVCVCCV	63	1,01
CCV	59	0,95

Tabulka 16. Nejfrekventovanější CV vzorce ve zpracovávaných řečových projevech spolu s jejich procentuálním podílem na všech taktech v materiálu. Tři nejčastější dvouslabičné vzorce jsou tučně.

### 3.1.3 Materiál z perspektivy experimentu

Pro účely experimentu prezentovaného v této práci bylo potřeba přihlédnout k faktu, že existuje velké množství kombinací konkrétních hlásek v jednotlivých vzorcích. Pro srovnatelnost výsledků bylo tedy třeba omezit vzorek vybraný pro samotný experiment, který měl zohlednit frekventovanost vzorců v projevech, aby se přiblížil skutečnému použití jazyka jakožto komunikačního prostředku. Byly proto vybrány tři v materiálu

nejčastější dvouslabičné konsonanticko-vokální vzorce, jejichž obsah byl dále podroben zkoumání.

Ze vzorku určeného pro experiment byly vyřazeny takty, které začínaly neznělou okluzivou, neboť u těchto jednotek často nebývá možné přesně určit, kdy začíná první hláska, což pro výzkum temporálních vlastností hlásek a mluvních taktů představuje komplikaci. Celkem bylo pro vzorek vybráno 338 taktů se vzorcem CVCV, 218 taktů se vzorcem CVCCV a 140 taktů odpovídající vzorci CCVCV. Do užšího výběru položek pro experiment se tak dostalo 696 mluvních taktů s různým hláskovým obsahem.

Po vydefinování těchto vzorců bylo následně zjištěno, co je hláskovým obsahem jednotlivých mluvních taktů. Vzorec CVCV obsahoval 142 typů hláskového obsahu, vzorec CVCCV 106 typů a vzorec CCVCV 84 typů. Nejfrekventovanější mluvní takty, které odpovídaly daným konsonanticko-vokálním vzorcům, jsou uvedeny v tabulce 17.

CVCV			CVCCV			CCVCV	
bude	21		dalji:	11		vla:di	11
budou	15		jedna	11		slabi:	6
roku	12		metru:	10		vla:da	5
nebo	11		mi:sti	10		sta:le	5
jiri:	11		mi:rni:	9		slabe:	4
mezi	10		jesce	9		vla:je	4
liji:	9		dobre:	8		snehu	4
zemi:	9		nejsou	6		v notsi	4
ra:no	7		zemje	5		sta:ti	4
nove:	6		dobri:	5		mjele	3
bili	6		ledna	5		vla:du	3
			notjni:	5		ftjera	3
						s toho	3
						znovu	3

Tabulka 17. Hláskové reprezentace jednotlivých fonotaktických vzorců, nejfrekventovanějších 10 typů od každého z nich. Transkripce odpovídá konvenci Mezinárodní fonetické abecedy (1999).

V rámci vzorců existovalo 332 typů hláskového obsahu a pro potřeby experimentu bylo nutné provést zobecnění. Jednotlivé takty tedy byly převedeny z konkrétních struktur na vzorce zohledňující přítomnost základního tónu, sonoritu a vokaličnost, tedy rysy, které jsou pro vnímání rytmu určující. Teoreticky by v rámci vzorců CVCV, CCVCV a CVCCV mohlo existovat 63 takových struktur, z nichž se ale v materiálu objevilo pouze 47. Nejvhodnější kandidáti do experimentu byli vybráni pomocí několika kritérií:

1. korektní výslovnost (vyloučeny byly zejména takty s desonorizovanými nebo elidovanými hláskami);
2. trvání celého taktu nesmělo přesáhnout průměrné trvání taktu v rámci dané skupiny o více než jednu směrodatnou odchylku; stejně tak položka nesměla trvat kratší dobu než průměrné trvání bez jedné směrodatné odchylky; skupinu taktů tvořil vzorec a kombinace samohlásek různých či stejných délek (např. všechny takty vzorce CVCV s první samohláskou krátkou a druhou dlouhou tvořily jednu skupinu, takty téhož vzorce s dvěma krátkými samohláskami tvořily skupinu druhou apod.);
3. nepřítomnost expresivní intonace a melodému; tj. preference taktů, na nichž se nerealizovalo melodické schéma vyšší prozodické jednotky.

Některé struktury byly zastoupeny silně, například vzorec mluvního taktu ZVSV (Z = znělý obstruent, V = vokál, S = sonora, N = neznělý obstruent) zahrnoval 28 různých hláskových realizací. V takových případech byly upřednostněny ty takty, které se v materiálu vyskytly nejčastěji (viz tabulka 18). Jiné vzorce měly zastoupení minimální, a pokud ani jedna realizace nesplňovala kritéria (příliš dlouhé/krátké trvání, desonorizace, vynechání hlásky, expresivní intonace), byl vzorec z testu vyřazen: takových vzorců bylo celkem 15. Vzorce, které se vyskytovaly v materiálu a které byly vybrány do experimentu, jsou zaznamenány v tabulce 18.

Za každý ze 32 zbývajících vzorců byl vybrán jeden, maximálně dva zástupci, splňující výše uvedená kritéria nejlépe. Tyto mluvní takty už přímo figurovaly v percepčním experimentu.

Na základě výše zmíněných rysů, které by mohly mít vliv na vnímání rytmu mluvního taktu, byly vzorce rozděleny do dvojic. Tyto dvojice struktur se vždy měly lišit jen jedním rysem (např. dvojice NVNV a NVZV se lišila pouze znělostí, dvojice ZVZV a ZZVZV počtem segmentů v iniciálním shluku apod.). Tyto dvojice vytvořily čtyři skupiny podle rysů, kterými se lišily: **přítomnost základního tónu v konsonantu** (9 dvojic), **počet segmentů v souhláskovém shluku** (10 dvojic), **pozice souhláskového shluku v taktu** (10 dvojic), **přítomnost formantového spektra v konsonantu** (12 dvojic). Aby tyto skupiny byly vzájemně porovnatelné co do počtu dvojic, přítomnost základního tónu a formantového spektra v souhlásce se rozlišovaly pouze ve skupině dvojic, které se odlišovaly výhradně touto vlastností; v ostatních skupinách (rozdíly v počtu segmentů



v konsonantické skupině a v jeho pozici) byly sonorní i znělé nesonorní hlásky považovány obecně za znělé (značené Z). Ve skupině párů lišících se přítomností základního tónu v konsonantu rozdíly v sonoritě nefigurovaly: tyto dvojice byly sestaveny pouze ze vzorců se znělými nesonorními a neznělými hláskami.

CVCV	CVCCV	CCVCV
ZVZV	ZVZZV	ZZVZV
NVNV	ZVNNV	ZZVNV
ZVNV	NVNNV	NNVNV
NVZV	NVZZV	NNVZV
ZVSV	SVSSV	NZVZV
NVSV	SVSNV	SSVSV
SVNV	SVNNV	SSVZV
SVZV	SVSZV	ZSVZV
SVSV	SVZZV	NSVZV
	ZVSZV	ZSVNV
	NVSZV	NSVSV
	NVSNV	NNVSV
	ZVSNV	ZSVSV
	NVSSV	ZZVSV
	NVNSV	NZVSV
	ZVSSV	SZVNV
	ZVZSV	
	ZVNSV	
	NVZSV	
	SVNZV	
	SVZSV	
	SVNSV	

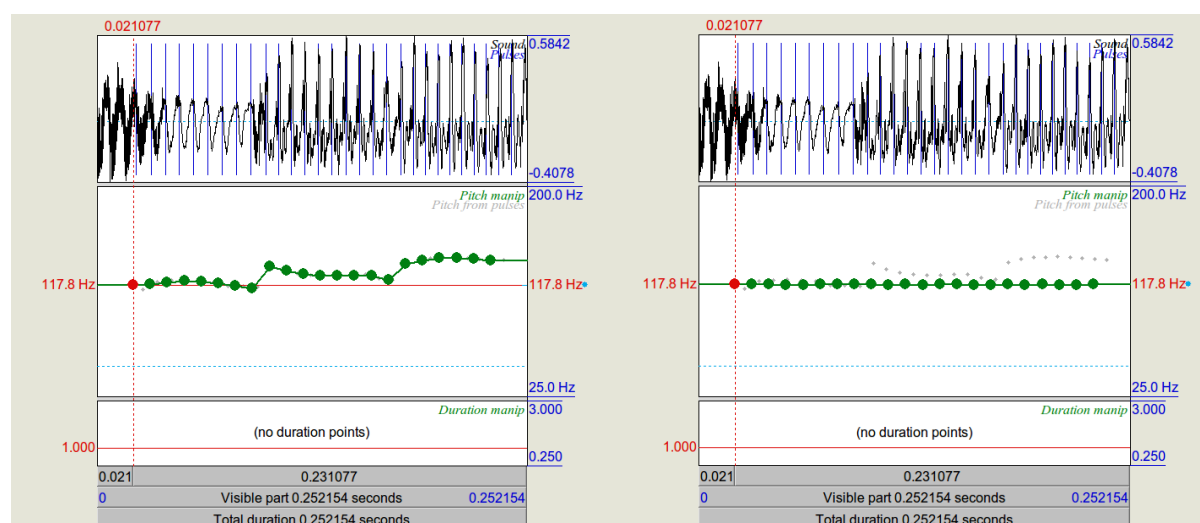
Tabulka 18. Struktury jednotlivých CV vzorců, které byly zastoupeny v materiálu, se zohledněním přítomnosti/absence základního tónu, sonority a vokaličnosti. Z = znělý obstruent, N = neznělý obstruent, V = vokál, S = sonora. Vzorce, jejichž realizace nesplňovaly kritéria pro experiment, jsou šedě.

## 3.2 Popis percepčního testu

### 3.2.1 Položky

Položka v percepčním testu sestávala ze dvou přirozeně proslovených dvouslabičných mluvních taktů nebo z přirozeného mluvního taktu a nízkofrekvenčního obrazu druhého taktu. Nízkofrekvenční obraz slova byl získán pomocí nízkopropustného filtru, nastaveného shodně u všech taktů na 400 Hz. Toto opatření bylo přijato proto, aby se posluchač maximálně odpoutal od segmentálního obsahu slova a vnímal pouze jeho předpokládané rytmické činitele. Filtrovaným položkám pak byla zvýšena intenzita o 5 dB pro dosažení percepční srovnatelnosti – vlivem rozsahu sluchového pole byly totiž položky s vyššími frekvencemi vnímány jako hlasitější (Palková, 1994). Některé

položky, obsahující jak pouze takty, tak i nízkofrekvenční obrazy, byly duplikovány a byla u nich pomocí techniky PSOLA (Pitch Synchronous Overlap and Add) manuálně vyhlazena intonace (obr. 3), aniž by však tato manipulace zásadně narušila vnímání přirozenosti řeči. Také tyto položky byly vybrány tak, aby reprezentovaly všechny zkoumané rysy. Položky s vyhlazenou intonací byly přidány pro případné ověření dříve zjištěných poznatků, že proměnlivá či vyšší  $f_0$  přispívá k vnímání delšího trvání zvuku než u stejných úseků s vyrovnanou základní frekvencí (Donovan & Darwin, 1979; Brigner, 1988; Cumming, 2011; Šimko *et al.*, 2015; Dawson *et al.*, 2017).



Obr. 3. Mluvní takt [zboru] s přirozenou a vyhlazenou intonací.

V rámci jedné položky byla mezi takty nebo taktem a obrazem taktu tichá pauza trvající 750 ms; po druhém taktu nebo obrazu taktu byla opět tichá pauza trvající 1,5 s. Poté následovala desenzitační pasáž (přehrání kytarového akordu pozpátku) s tichou pauzou (1,5 s). Jedna položka trvala celkem cca 7 s.

Celkem 210 položek bylo pseudonáhodně seřazeno a následně rozděleno do tří bloků po 70 položkách s následujícími variantami:

varianta položky	počet položek
slovo X + slovo Y	41
slovo Y + slovo X	41
slovo X + nízkofrekvenční obraz slova Y	41
slovo Y + nízkofrekvenční obraz slova X	41
slovo X + slovo Y (vše s vyhlazenou intonací)	5
slovo Y + slovo X (vše s vyhlazenou intonací)	5
slovo X + nízkofrekvenční obraz slova Y (vše s vyhlazenou intonací)	5
slovo Y + nízkofrekvenční obraz slova X (vše s vyhlazenou intonací)	5

slovo X + nízkofrekvenční obraz slova X / slovo Y + nízkofrekvenční obraz slova Y	10
slovo X + nízkofrekvenční obraz slova X / slovo Y + nízkofrekvenční obraz slova Y (vše s vyhlazenou intonací)	3
opakované položky	13
celkem	210

Tabulka 19. Typy položek percepčního testu a jejich počet.

Opakované položky, které v testu figurovaly kvůli sledování konzistentnosti respondentů, byly stejnoměrně vybrány z jednotlivých skupin vzorců a byly pseudonáhodně zařazeny do testu. Položky, v nichž se vyskytoval mluvní takt spolu se svým vlastním nízkofrekvenčním obrazem, byly obměňovány s každou ze tří verzí testu.

### 3.2.2 Respondenti

Po vybrání definitivního počtu položek byly tyto třikrát pseudonáhodně seřazeny do jednotlivých variant percepčního testu. Testy pak byly spolu s formulářem jednotným pro všechny varianty předány respondentům. Všichni respondenti byli rodilými mluvčími češtiny ve věku 18–65 let. Ze 40 osob bylo 25 mužů a 15 žen. Věkové rozpětí respondentů je rozepsáno v tabulce 20.

věk	respondentů
18–20 let	1
21–30 let	17
31–40 let	13
41–50 let	5
>50 let	2
neuvedeno	2
celkem	40

Tabulka 20. Rozdělení respondentů do skupin podle věku.

místo narození	respondentů
Praha	12
Čechy	21
Morava	2
Slezsko a MS pomezí	4
narozen/a mimo ČR	1
celkem	40

Tabulka 21. Rozdělení respondentů do skupin podle místa narození.

Z tabulek je patrné, že většina respondentů byla mezi 20–40 lety věku (30 osob ze 40) a že pokusné osoby pocházely převážně z Prahy a Čech (33 osob ze 40). Města na Moravě a ve Slezsku označilo jako své místo narození 15 % respondentů.

### 3.2.3 Instrukce a průběh experimentu

Respondenti dostali přístup ke zvukovým souborům, nahraným do cloudového úložiště. Zvuky obsahovaly tři zácvičné položky a jednotlivé osmiminutové bloky testu po 70 položkách. Všechny pokusné osoby obdržely identické psané instrukce, v nichž bylo uvedeno, že úkolem respondenta je posoudit, nakolik jsou si zvuky ve dvojici rytmicky

podobné. Rytmická podobnost byla blíže specifikována pouze v případě, že respondent výslovně žádal upřesnění; obecně však odpovědi pokusných osob měly vystihovat intuitivní interpretaci rytmu jako ze značné části percepčního jevu. Podobnost respondenti určovali na škále 1–5, kde 1 = téměř nebo zcela identické, 5 = rytmicky výrazně odlišné. Instrukce obsahovaly také dovětek, že by se hodnotící osoba měla pokusit odpoutat od melodie a významů slov v jednotlivých položkách a hodnotit pouze rytmus (dle intuitivní interpretace). Respondenti byli dále instruováni, aby test absolvovali v tichém prostředí a s kvalitními sluchátky a aby si mezi jednotlivými bloky testu udělali krátkou přestávku.

Formulář, který pokusné osoby vyplňovaly, byl distribuován jako dotazník na službě Google Docs a hodnocení položky na škále respondenti zaklikávali vždy u čísla položky.

### **3.2.4 Statistická analýza dat**

V rámci obecné analýzy byl zkoumán vztah průměrného hodnocení položek s některým z rysů, které mohly mít vliv na vnímání rytmické podobnosti nebo odlišnosti porovnávané dvojice. Hodnocení položky bylo v analýzách považováno za závislou proměnnou a přítomnost nebo nepřítomnost rysu za nezávislou kategorickou nebo řadovou proměnnou. Pro ověření trendů byly využity t-testy pro nezávislé výběry dle skupin nebo výpočet korelace mezi hodnocením jakožto závislou proměnnou a rysy reprezentovanými spojitými numerickými proměnnými. Všechny souhrnné analýzy byly provedeny dvakrát, a to jak na všech položkách testu, tak pouze na položkách, které neobsahovaly slovo a nízkofrekvenční obraz téhož slova prosloveného týměž mluvčím.

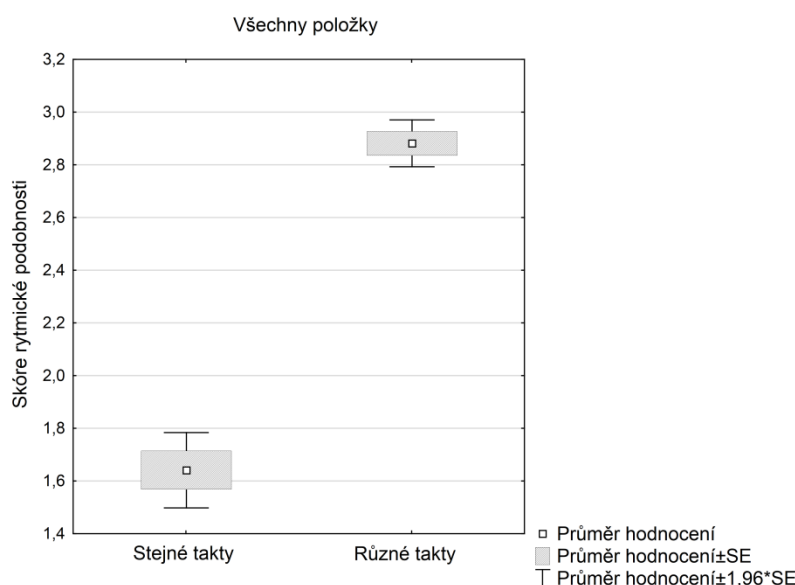
## 4 Výsledky

### 4.1 Vlastnosti experimentálního plánu

#### 4.1.1 Shoda nízkofrekvenčního obrazu s původním taktem

Některé z položek obsahovaly mluvní takt a jeho vlastní nízkofrekvenční obraz (viz oddíl 3.2.1). Tyto položky byly do testu zařazeny zejména proto, aby posluchačům poskytly vztažný bod rytmické podobnosti. Jelikož byly tyto položky specifické svou rytmickou identičností i totožností mluvního, byly všechny obecné analýzy v této kapitole provedeny jak na hodnocení celé sady položek, tak na hodnocení souboru položek, který neobsahoval rytmicky identické takty v rámci jedné položky.

Vliv shodných taktů v položce prokázal t-test:  $t(221) = 11,9$ ;  $p < 0,001$ . Zatímco hodnocení položek, které obsahovaly různé takty (ať už v plné formě, nebo jako nízkofrekvenční obraz), se pohybovalo kolem průměru 2,88, hodnocení rytmicky identických položek bylo průměrně 1,64.

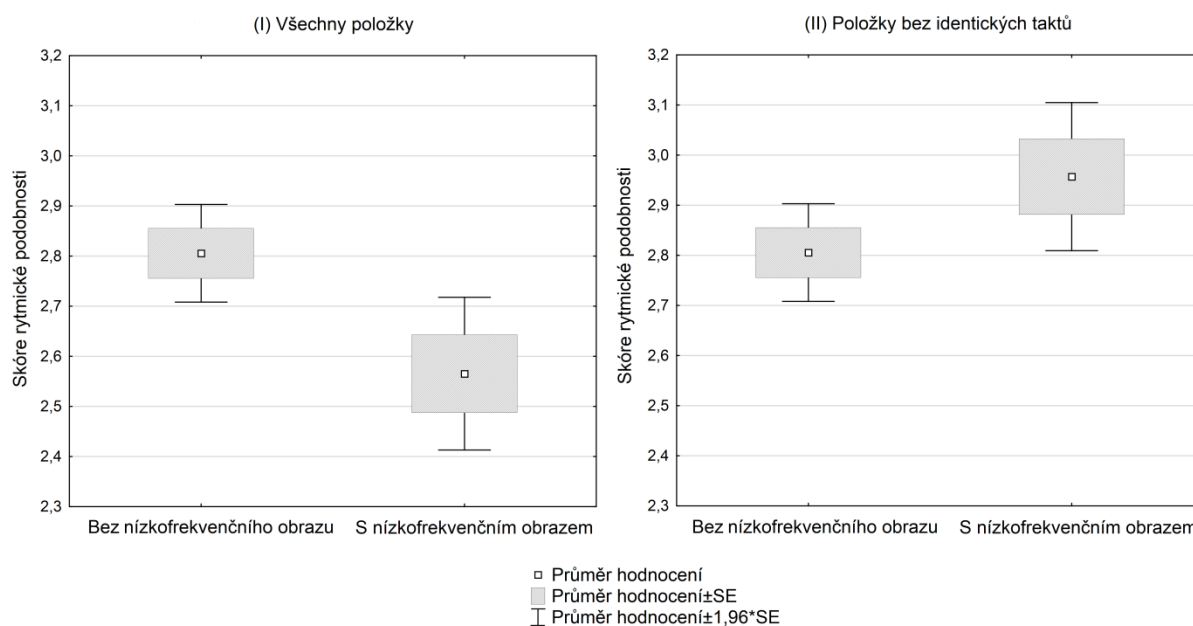


Graf 3. Krabicový graf závislosti průměrného hodnocení položky na přítomnosti taktu a jeho vlastního nízkofrekvenčního obrazu v položce. Zkratka SE značí směrodatnou chybu; skóre rytmické podobnosti na ose y: 1 = velmi podobné, 5 = velmi rozdílné.

#### 4.1.2 Přítomnost nízkofrekvenčního obrazu v položce

Dalším faktorem v hodnocení položek byla samotná skladba položky, která mohla obsahovat buď dva různé mluvní takty i s hláskovým obsahem (např. [bɪlɪ] a [bude]), nebo jeden takt a jeden nízkofrekvenční obraz jiného či stejného taktu (např. takt [mezi] a obraz taktu [ɲimi]). T-testy ukázaly, že posluchači vnímali jako rytmicky podobnější

položky, v nichž se vyskytoval nízkofrekvenční obraz taktu. Tento vztah se ale projevil pouze tehdy, když byly do analýzy zahrnuty i položky s takty a jejich vlastními nízkofrekvenčními obrazy ( $t(221) = -2,36; p < 0,05$ ); v případě vynechání těchto položek byl zaznamenán nevýznamný trend opačným směrem ( $t(182) = 1,68; p \doteq 0,095$ ) – je tedy otázkou, jestli vnímání rytmu skutečně ovlivnila přítomnost, či absence segmentálního obsahu druhého taktu, nebo obecné hodnocení shodných položek jako výrazně rytmicky obdobnějších vzhledem k ostatním.

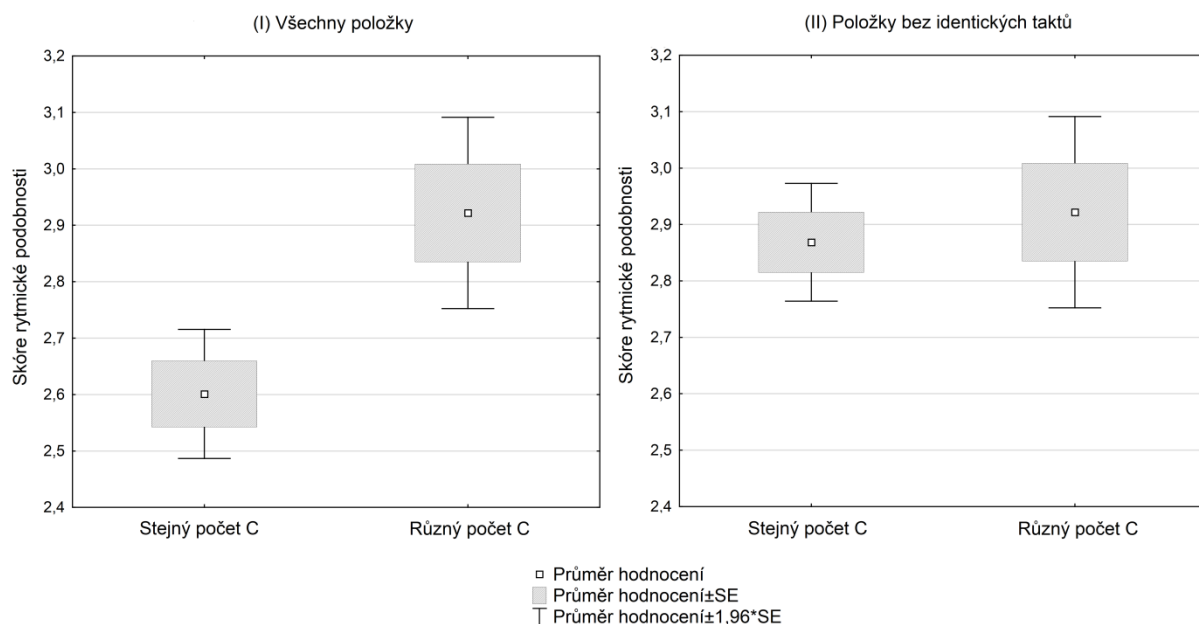


Graf 4 (I, II). Krabicový graf závislosti průměrného hodnocení položky na přítomnosti nízkofrekvenčního obrazu taktu v položce. V grafu I jsou zahrnuty i položky, které obsahují takt s jeho vlastním nízkofrekvenčním obrazem, graf II postihuje soubor položek bez rytmicky identických taktů.

## 4.2 Fonotaktické vlastnosti a vnímání rytmu

### 4.2.1 Počet konsonantů v souhláskové skupině

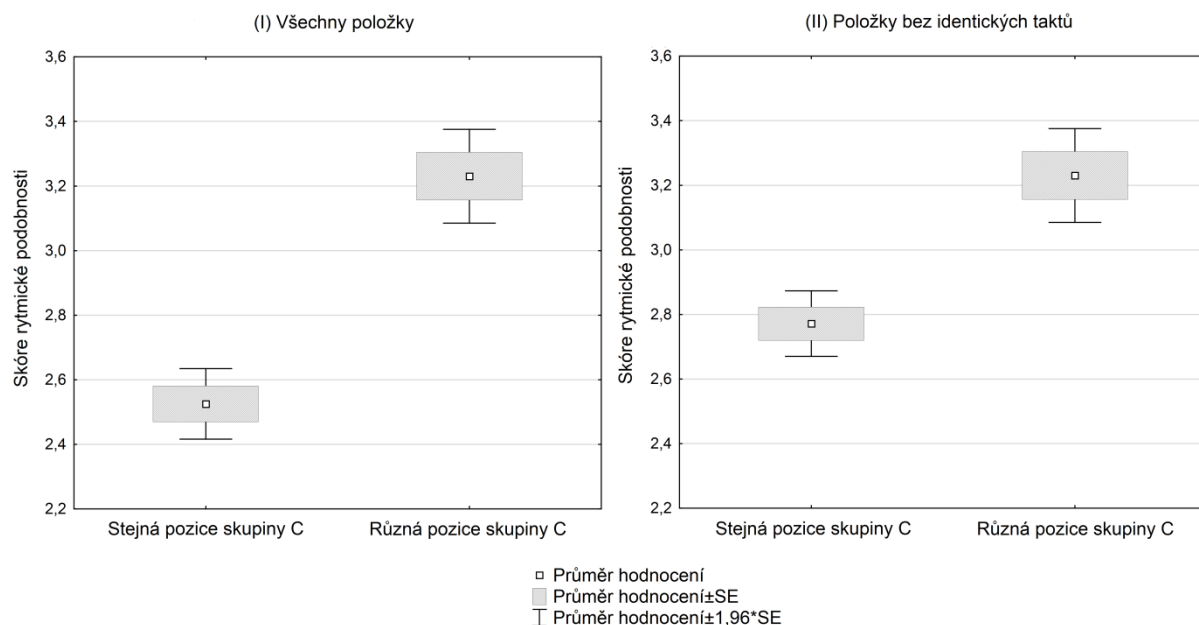
Výsledky T-testů položek, které se lišily pouze počtem konsonantů, oproti ostatním ukázaly statisticky významný vliv rozdílu v počtu konsonantů v souhláskové skupině na vnímání rytmické podobnosti. Takty v položce, které měly rozdílný počet souhlásek ve shluku (např. dvojice NVZV–NNVZV [fa:zɛ]–[sta:lɛ]), byly posluchači průměrně hodnoceny jako rytmicky vzdálenější než položky, které tuto odlišnost neobsahovaly. Rozdíl v hodnocení byl robustnější, pokud byly analyzovány všechny položky, tedy i ty, které obsahovaly takt s jeho vlastním nízkofrekvenčním obrazem:  $t(221) = -2,56; p < 0,05$ ; v případě vynechání těchto položek byl patrný pouze nevýznamný trend podobného směru:  $t(182) = -0,5; p \doteq 0,617$ .



Graf 5 (I, II). Krabicový graf závislosti průměrného hodnocení položky na přítomnosti rozdílu v počtu konsonantů v souhláskové skupině. V grafu I jsou zahrnuty i položky, které obsahují takt s jeho vlastním nízkofrekvenčním obrazem, graf II postihuje soubor položek bez rytmicky identických taktů.

#### 4.2.2 Pozice souhláskového shluku v taktu

Položky, jejichž takty se odlišovaly pozicí souhláskové skupiny v taktu (např. dvojice ZVZZV–ZZVZV [volb<sub>1</sub>]–[vj<sub>1</sub>l<sub>1</sub>]), byly hodnoceny jako méně si rytmicky podobné než ostatní položky, v nichž se tento rozdíl neobjevoval.

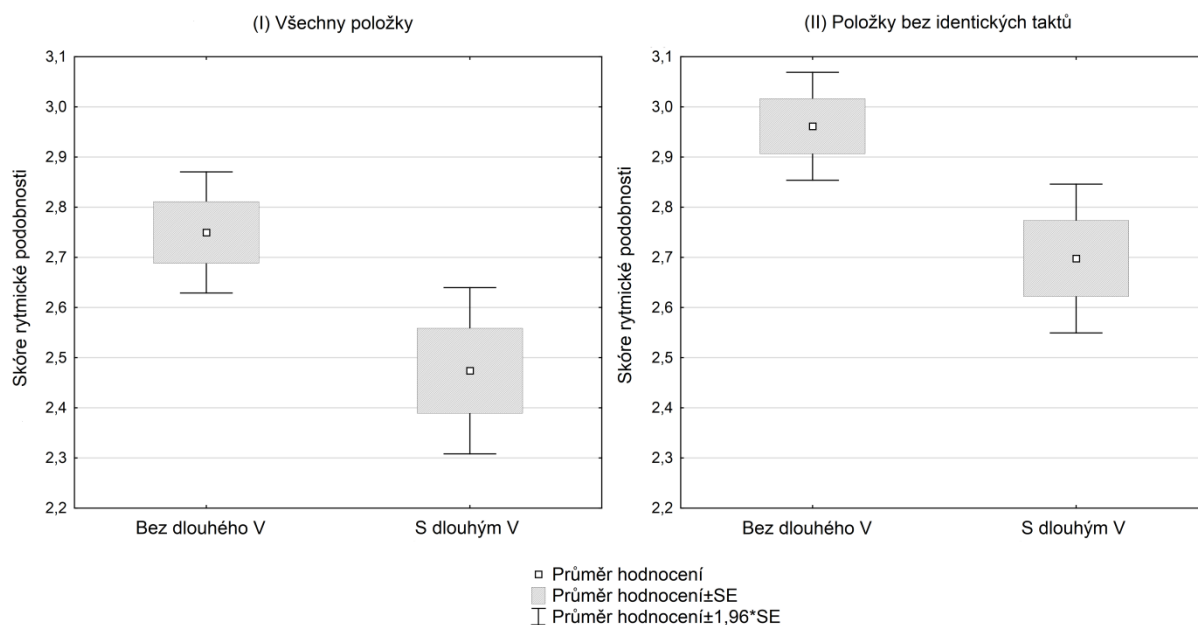


Graf 6 (I, II). Krabicový graf závislosti průměrného hodnocení položky na přítomnosti rozdílu v pozici konsonantů v taktu. V grafu I jsou zahrnuty i položky, které obsahují takt s jeho vlastním nízkofrekvenčním obrazem, graf II postihuje soubor položek bez rytmicky identických taktů.

Tento vliv byl patrný jak v analýze všech položek ( $t(221) = -5,96; p < 0,001$ ), tak při započítání pouze těch položek, které neobsahovaly mluvní takt s jeho vlastním nízkofrekvenčním obrazem:  $t(182) = -4,54; p < 0,001$ . Ve srovnání s předchozí proměnnou, tedy počtem konsonantů v souhláskové skupině, je efekt pozice souhláskového shluku výraznější.

#### 4.2.3 Vliv dlouhé samohlásky v taktu

Další zkoumanou vlastností s předpokládaným vlivem na vnímání podobnosti či odlišnosti rytmu mluvních taktů byla přítomnost dlouhého vokálu v porovnávaných taktech. Pokud se v jedné testové položce vyskytoval dlouhý vokál, byl vždy v obou srovnávaných mluvních taktech ve stejné slabice (porovnávaly se tedy například takty [sta:ti] a [svɛ:ɦo] nebo [budɛ] a [ɲimi]), neboť v případě, kdy by v jednom z porovnávaných taktů dlouhý vokál byl, a ve druhém ne, by bylo obtížné vysledovat jiný vliv na vnímání rytmu než právě rozdíl v přítomnosti dlouhé samohlásky – lze například předpokládat, že kdybychom srovnávali v rámci jedné položky takty [sta:ti] a [ɲimi] nebo [svɛ:ɦo] a [zvɛda:], byla by pozornost posluchačů upřena na rozdíl v délce samohlásky, a subtilnější rozdíly by proto mohly být při vnímání rytmu potlačeny.



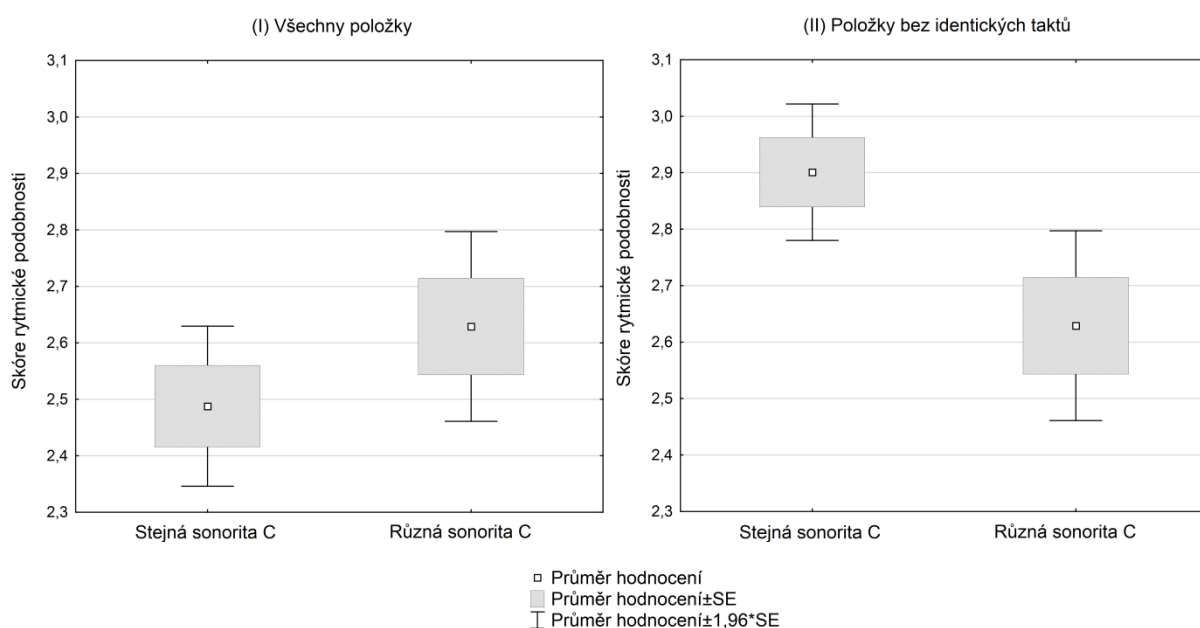
Graf 7 (I, II). Krabicový graf závislosti průměrného hodnocení položky na přítomnosti dlouhé samohlásky v taktech položky. V grafu I jsou zahrnuty i položky, které obsahují takt s jeho vlastním nízkofrekvenčním obrazem, graf II postihuje soubor položek bez rytmicky identických taktů.



Podle výsledku t-testu byly mluvní takty, které obsahovaly pouze krátké vokály, hodnoceny posluchači jako méně si rytmicky podobné než takty, jež dlouhou samohlásku obsahovaly, což se potvrdilo jak v souhrnné analýze všech položek ( $t(221) = -2,55$ ;  $p < 0,05$ ), tak v rámci testování položek bez zahrnutí položek s identickými takty:  $t(182) = -2,72$ ;  $p < 0,01$ . Položky s dlouhými vokály také vykazovaly menší rozdíly v celkovém trvání srovnávaných taktů. Oproti předchozí proměnné byl vliv dlouhé samohlásky slabší, nicméně byl srovnatelný s působením počtu konsonantů ve skupině (všechny položky) nebo silnější (položky bez identických taktů).

#### 4.2.4 Rozdíl v sonoritě znělých souhlásek

Souhrnná analýza vlivu rozdílu v sonoritě souhlásek ukázala neprůkazný trend, kdy položky obsahující znělé souhlásky odlišné sonority (např. ZZVSV a ZSVSV [zboru] a [vjelɪ]) byly hodnoceny jako méně si rytmicky podobné než ostatní položky:  $t(173) = -1,17$ ;  $p = 0,244$ . Při analýze neobsahující položky s identickými takty se objevil trend opačný – položky lišící se v sonoritě znělých konsonantů byly tedy hodnoceny jako rytmicky méně rozdílné než jiné:  $t(134) = 2,64$ ;  $p < 0,01$ .



Graf 8 (I, II). Krabicový graf závislosti průměrného hodnocení položky na rozdílu v sonoritě souhlásek v rámci jedné položky. V grafu I jsou zahrnuty i položky, které obsahují takt s jeho vlastním nízkofrekvenčním obrazem, graf II postihuje soubor položek bez rytmicky identických taktů.

Může se zdát, že tento výsledek jde proti intuici, nicméně pokud přihlídneme k designu experimentu, je vysvětlení poměrně jednoduché: po odstranění položek se stejnými

mluvními takty zůstaly k porovnání proti vlivu sonority pouze položky s rozdíly v jiných vlastnostech, z nichž některé lze považovat za silnější než subtilní rozdíly v sonoritě souhlásek (např. počet konsonantů ve skupině, pozice souhláskového shluku). Pro zkoumání vlivu rozdílu sonority a znělosti by tedy bylo třeba porovnávat pouze soubory položek s rozdílem sonority se souborem položek se stejnou sonoritou a bez dalších fonotaktických odlišností; taková analýza však přesahuje rámec této práce, jejímž cílem je vystihnout nejsilnější vlivy fonotaktického uspořádání českých mluvních taktů na vnímání rytmu rodilými mluvčími.

#### 4.2.5 Kontrast znělosti

Analýza vnímání rytmické podobnosti dále neprokázala významný vliv odlišné znělosti u znělých nesonorních a neznělých souhlásek v rámci jedné položky. Příkladem takových položek byly dvojice NVNV–ZVNV [sôu to]–[ʒa:tsɪ] či ZVZV–ZVNV [mɛzi]–[jako]. Trendy byly neprůkazné jak u analýzy všech položek ( $t(221) = 1,31$ ;  $p \doteq 0,192$ ), tak bez zahrnutí položek s identickými takty ( $t(182) = -0,88$ ;  $p \doteq 0,381$ ). I zde, stejně jako u neprůkazného rozdílu ve vnímání taktů s rozdílnou sonoritou souhlásek, platí, že vynesení jednoznačného závěru by měl předcházet experiment zaměřující se výhradně na tento rys.

Je vhodné poznamenat, že fonologická znělost může být signalizována i jinými faktory, než je přítomnost základní frekvence. U párových konsonantů bývá trvání neznělých hlásek delší než trvání hlásek znělých; v některých jazycích (např. v angličtině) se také projevují kompenzační strategie, co se týče trvání samohlásky předcházející znělou, či neznělou souhlásku: pokud následuje znělý konsonant, bývá vokál před ním delší než samohláska, za níž se vyskytuje neznělý obstruent (Kent & Read, 2002). Ačkoli vliv temporálních kompenzací v češtině nebyl systematicky zkoumán, výsledky Borovičkové a Maláče (1967) a Machače (2006) jisté tendence v podobném směru naznačují. Silnějším vodítkem v určení fonologické znělosti se ukázala koncová část vokálu předcházejícího konsonant (např. Hogan & Rozsypal, 1980). Mluvčí mohou naznačit znělostní kontrast přítomností základního tónu během závěru, trváním závěru, silou exploze nebo aspirace (např. v angličtině), ale i frekvencí základního tónu a prvního formantu: Výsledky Hogana a Rozsypala (1980) nebo Castlemanové a Diehla (1996) odhalily, že základní frekvence a první formant vokálu na předělu samohláska-souhláska byly nižší, pokud za samohláskou následovala závěrová hláska znělá, oproti případu, kdy

vokál předcházet okluzivu neznělou. Všechny tyto rysy mohou být posluchačem vnímány a používány, kdykoli je to zapotřebí, ale je rovněž možné, že jednotlivé rysy přispívající k vnímání fonologické znělosti tvoří hierarchii, v níž jsou temporální aspekty vnímány, jen pokud ostatní vodítka k určení znělosti absentují nebo jsou nejednoznačná (Kent & Read, 2002).

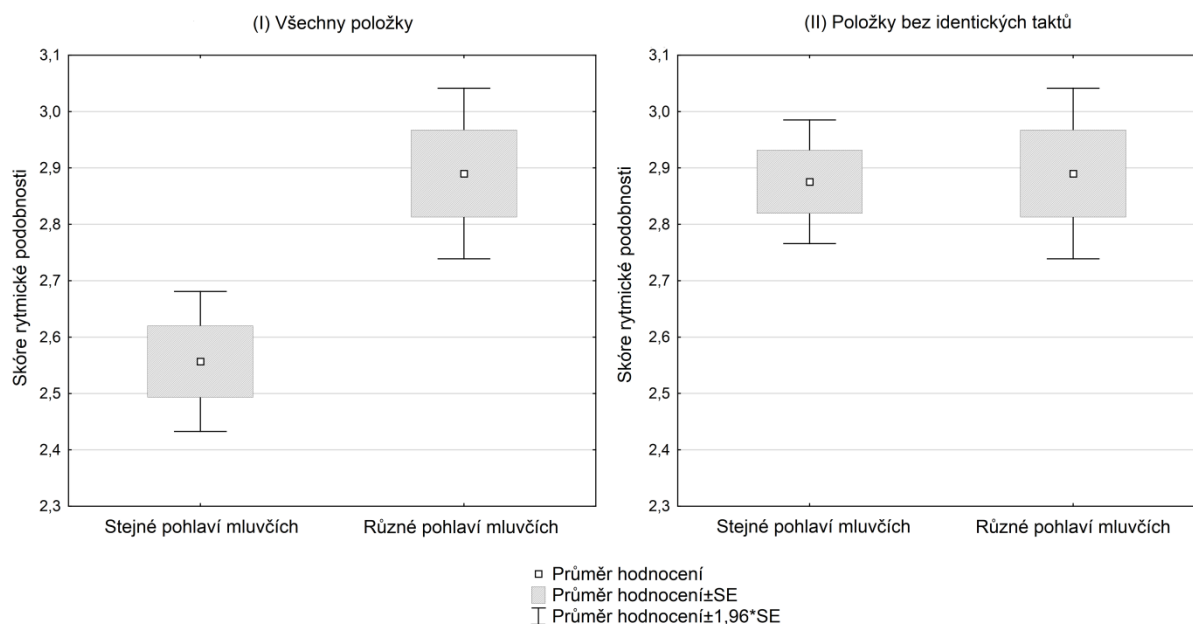
### **4.3 Suprasegmentální a biologické faktory**

#### **4.3.1 Vyhlazení intonačního průběhu**

Všechny takty v položkách percepčního testu byly vybírány tak, aby neobsahovaly výrazné intonační rozdíly. Pro ověření předpokladu, že by změny melodie neměly zasahovat do hodnocení podobnosti rytmu taktů v rámci položky, byly některé z položek do testu zařazeny několikrát, a to jednak bez umělých zásahů do intonačního průběhu, jednak s vyhlazením  $f_0$  do homogenní linie. T-testy neodhalily žádný průkazný vliv vyhlazené, či původní intonační linie na percepci taktů jako rytmicky si podobných, nebo odlišných. Analýza provedená na všech položkách ukázala pouze marginální významnost proměnné ( $t(221) = 1,91$ ;  $p \doteq 0,057$ ), kdy vyhlazené položky byly v průměru hodnoceny jako rytmicky si podobnější než ostatní. Dále byl analyzován soubor položek s vyloučením identických taktů ( $t(182) = 1,51$ ;  $p \doteq 0,133$ ) a také pouze položky obsahující homogenní intonační linii spolu s jejich původními protějšky, opět se zahrnutím i s vyloučením položek s týmiž takty ( $t(52) = -0,23$ ;  $p \doteq 0,822$ ;  $t(38) = 0,01$ ;  $p \doteq 0,99$ ).

#### **4.3.2 Rozdíl v pohlaví mluvčích v rámci jedné položky**

Při zkoumání vlivu odlišného pohlaví mluvčích obou taktů proslovených v rámci jedné položky se ukázal statisticky významný rozdíl pouze v souhrnné analýze všech položek:  $t(221) = 3,14$ ;  $p < 0,01$ , přičemž lépe hodnoceny byly položky, u nichž se pohlaví mluvčích shodovalo. Při provedení t-testu pouze na položkách bez identických taktů vyšel trend stejného směru, ale statisticky zcela nevýznamný:  $t(182) = 0,16$ ;  $p \doteq 0,876$ . Z těchto výsledků je možné usuzovat, že vnímaný rozdíl spíše ukazoval na vliv individuality mluvčího než pohlaví jako faktoru pro vnímání rytmické podobnosti.



Graf 9 (I, II). Krabicový graf závislosti průměrného hodnocení položky na rozdíl pohlaví mluvčích taktů v rámci jedné položky. V grafu I jsou zahrnuty i položky, které obsahují takt s jeho vlastním nízkofrekvenčním obrazem, graf II postihuje soubor položek bez rytmicky identických taktů.

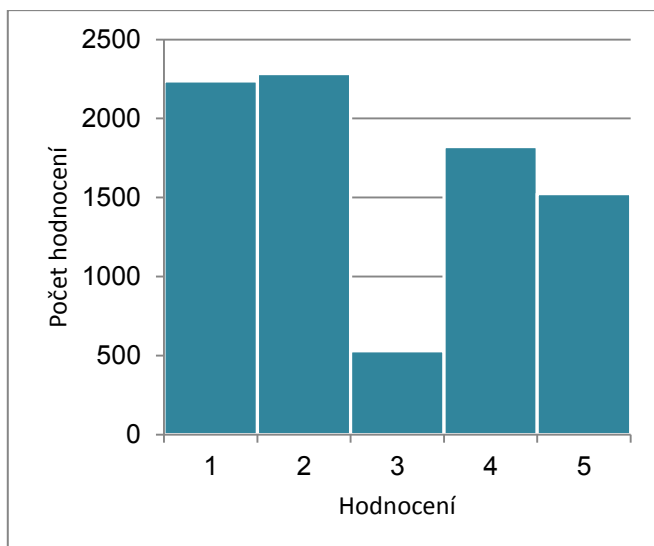
rys v rámci položky   síla efektu	t (všechny položky)	t (položky s vyloučením identických taktů)
shoda nízkofrekvenčního obrazu	<b>11,9</b>	
přítomnost nízkofrekvenčního obrazu	<b>-2,36</b>	1,68
počet konsonantů ve shluku	<b>-2,56</b>	0,5
pozice konsonantické skupiny v taktu	<b>-5,96</b>	<b>-4,54</b>
dlouhá samohláska	<b>-2,55</b>	<b>-2,72</b>
rozdíl v pohlaví mluvčích	<b>3,14</b>	0,16
kontrast sonority	-1,17	<b>2,64</b>
kontrast znělosti	1,31	-0,88
vyhlazení intonace	1,91	1,51

Tabulka 22. Porovnání síly jednotlivých vlivů na vnímání rytmické podobnosti pomocí srovnání hodnot testového kritéria *t*. Statisticky významné vlivy jsou tučně.

## 4.4 Položková analýza

### 4.4.1 Obecné trendy v hodnocení položek

Druhá sada analýz v této práci spočívá v prozkoumání těch položek, které v rámci srovnání oproti ostatním vykazovaly extrémní nebo nestandardní hodnoty v hodnocení. Aby bylo možné definovat, co v tomto případě znamená „extrémní“ či „nestandardní“, bylo nutné prozkoumat rozložení číselných hodnocení všech položek.



Graf 10. Sloupcový graf popisující rozdělení jednotlivých hodnocení všech položek v rámci celého testu.

hodnocení	počet hodnocení
1	2238
2	2285
3	531
4	1822
5	1524
celkem	8400

Tabulka 23. Počty jednotlivých typů hodnocení všech položek v testu.

Z grafu 10 a tabulky 23 je možné vyčíst nápadně menší frekventovanost prostředního hodnocení, tedy číselný ekvivalent vyjadřující „vstupy v položce neznějí rytmicky ani příliš podobně ani příliš odlišně“. Je tedy možné předpokládat, že i přes vysoký počet položek v rámci jednoho testu se respondenti jen zřídka utíkali k neutrálnímu hodnocení a zpravidla se přikláněli k jedné, nebo druhé možnosti, tedy hodnotili vstupy v položce spíše jako podobné, nebo odlišné.

Počty hodnocení byly následně pomocí testu chí-kvadrát porovnány s jejich rovnoměrným rozložením jakožto s očekávanými četnostmi hodnocení – nulová hypotéza tedy počítala s tím, že mezi jednotlivými hodnoceními co do počtu nebudou žádné rozdíly a každé hodnocení se v testu vyskytne ve 20 % z celkového počtu hodnocení.

počty hodnocení	1	2	3	4	5
$n_{\text{poz}}$	2238	2285	531	1822	1524
$n_{\text{oč}}$	1680	1680	1680	1680	1680

Tabulka 24. Pozorované ( $n_{\text{poz}}$ ) a očekávané ( $n_{\text{oč}}$ ) četnosti výskytu rozdílných hodnocení v testu.

Po stanovení očekávaných četností hodnocení byla vypočítána hodnota  $\chi^2$  ( $4; n = 8400$ ) = 1215,53;  $p < 0,001$ . Statistická významnost na hladině  $\alpha = 0,01$  byla potvrzena a nulová hypotéza o rovnoměrném rozdělení hodnocení v testu mohla být zamítnuta.

Nerovnoměrné zastoupení počtu hodnocení všech položek je tedy možné považovat za nenáhodné.

#### 4.4.2 Kritéria výběru extrémně hodnocených položek

Jelikož pro kategorické proměnné, jako je hodnocení na škále, není vhodné používat jako ukazatel střední hodnoty průměr a jako ukazatel variability směrodatnou odchylku, byly pro stanovení extrémně hodnocených položek využity ukazatele střední hodnoty medián a modus. Za uživateli nestandardně či extrémně hodnocené byly považovány položky následujících typů:

- 1) položky, které vykazovaly hodnotu mediánu hodnocení 1–1,5 nebo 4,5–5;
- 2) položky, jejichž hodnocení mělo více modů, které spolu nesousedily (např. byly v součtu hodnoceny nejčastěji jako 2 – relativně rytmicky podobné – a 5, tedy rytmicky výrazně či zcela odlišné) a z nichž alespoň jeden odpovídal okraji škály hodnocení.

Z položek splňujících první část prvního požadavku, tedy těch, jejichž medián se pohyboval v rozmezí od 1 do 1,5, byly vyloučeny položky obsahující slovo a jeho vlastní obraz. U těchto položek se totiž nikdy nevyskytoval medián či modus hodnocení vyšší než 2 (alespoň částečná rytmická podobnost). Již bylo prokázáno, že posluchači hodnotili tyto položky jako významně rytmicky si podobnější než ostatní položky. Z 39 takovýchto položek jich 31 vykazovalo medián hodnocení 1, u jedné položky byl zaznamenán medián 1,5 a medián zbylých sedmi položek byl roven dvěma – tyto položky byly také podrobeny individuální analýze.

#### 4.4.3 Vybrané položky

Na základě výše uvedených kritérií byly vybrány položky, jejichž hodnocení odpovídalo krajním hodnotám škály, na níž posluchači určovali rytmickou podobnost taktů v položce. Podrobný přehled těchto položek je uveden v tabulce 25.

pol.	takt 1	takt 2	počet hodn.	medián	modus	trvání 1 (ms)	trvání 2 (ms)	rozdíl trvání (ms)	rozdíl trvání %
P2	státy	<i>svého</i>	40	<b>4,5</b>	5	313,61	288,16	25,45	8,12
P4	se zde	<i>Šopko</i>	40	<b>1,5</b>	1	301,09	313,74	12,65	4,2
P10	bude	<i>v Gaze</i>	40	<b>5</b>	5	248,98	362,04	113,06	45,41
P11	jsou to	<i>státy</i>	40	<b>1,5</b>	1	287,12	313,61	26,49	9,23
P23	sešlo	<i>škody</i>	40	<b>4,5</b>	5	330,16	354,2	24,04	7,28

P27	sboru	<i>jedna</i>	40	2	<b>více</b>	252,15	303,45	51,3	20,35
P30	bude	<i>nimi</i>	40	5	5	248,98	219,86	29,12	11,7
P33	vládě	<i>jméno</i>	40	2,5	<b>více</b>	338,23	330,88	7,35	2,17
P34	v noci	<i>Lhase</i>	40	<b>4,5</b>	5	307,89	337,82	29,93	9,72
P35	výšce	<i>místy</i>	40	<b>1</b>	1	370,02	347,62	22,4	6,05
P56	země	<i>nimi</i>	40	5	5	332,74	219,86	112,88	33,92
P63	době	<i>sboru</i>	40	<b>4,5</b>	5	288,12	252,15	35,97	12,48
P66	jedna	<i>vjely</i>	40	<b>4,5</b>	5	303,45	238,78	64,67	21,31
P68	jedna	<i>sboru</i>	40	5	5	303,45	252,15	51,3	16,91
P71	nimi	<i>bude</i>	40	<b>1</b>	1	219,86	248,98	29,12	13,24
P73	sboru	<i>vjely</i>	40	<b>1,5</b>	1	252,15	238,78	13,37	5,3
P76	místy	<i>výšce</i>	40	<b>1</b>	1	347,62	370,02	22,4	6,44
P80	byli	<i>bude</i>	40	<b>1</b>	1	220,95	248,98	28,03	12,69
P82	mezi	<i>nimi</i>	40	5	5	251,07	219,86	31,21	12,43
P86	se zde	Šopko	40	<b>1</b>	1	301,09	313,74	12,65	4,2
P93	jsou to	<i>státy</i>	40	3,5	<b>více</b>	287,12	313,61	26,49	9,23
P169	výšce	<i>místy</i>	40	<b>1</b>	1	370,02	347,62	22,4	6,05
P173	jméno	<i>vládě</i>	80	<b>1</b>	1	330,88	338,23	7,35	2,22
P184	místy	<i>výšce</i>	40	<b>1,5</b>	1	347,62	370,02	22,4	6,44
N11	nimi	<i>nimi</i>	13	<b>2</b>	<b>více</b>	219,86	219,86	0	0
N16	sešlo	<i>sešlo</i>	14	<b>2</b>	2	330,16	330,16	0	0
N24	v Gaze	<i>v Gaze</i>	14	<b>2</b>	2	362,04	362,04	0	0
N26	země	<i>země</i>	14	<b>2</b>	<b>více</b>	332,74	332,74	0	0
N30	Lhase	<i>Lhase</i>	13	<b>2</b>	2	337,82	337,82	0	0
N32	silně	<i>silně</i>	13	<b>2</b>	2	303,31	303,31	0	0
N38	na dva	<i>na dva</i>	13	<b>2</b>	2	279,91	279,91	0	0

Tabulka 25. Rozpis položek vykazujících extrémní hodnocení. Ve sloupci *položky* jsou čísla analyzovaných položek. Položky s označením Px obsahují rozdílné takty nebo takt a nízkofrekvenční obraz jiného taktu, položky s označením Nx obsahují takt a nízkofrekvenční obraz téhož taktu. Ve sloupci *takt 1* je uveden slovní obsah prvního taktu položky; sloupec *takt 2* obsahuje obsah druhého taktu v položce. Text v tomto sloupci psaný kurzívou značí, že druhý takt v položce zazněl jen jako nízkofrekvenční obraz. V dalším sloupci je uveden celkový počet hodnocení dané položky a v následujících sloupcích medián a modus každé položky. Tučný text v těchto sloupcích značí kritérium výběru položky pro analýzu (např. položka P33 byla vybrána pro podrobnou analýzu kvůli tomu, že se v jejím hodnocení objevilo více modů, které spolu nesousedily a z nichž alespoň jeden dosahoval krajní hodnoty; položka P86 byla zkoumána proto, že medián jejího hodnocení byl roven jedné; položka N38 zase obsahovala takt s jeho vlastním obrazem, ale medián jejího hodnocení byl roven dvěma apod.) Poslední čtyři sloupce se zabývají trváním prvního a druhého taktu v položce, rozdílem jejich trvání v milisekundách a procentuálním rozdílem jejich trvání (za výchozí hodnotu 100 % se považovalo trvání prvního taktu).

#### 4.4.3.1 Položky s nízkým mediánem hodnocení

V rámci percepčního experimentu platilo, že čím nižší bylo hodnocení položky posluchači, tím podobnější co do rytmu respondentům oba mluvní takty v položce připadaly. Položky, které měly tedy medián hodnocení mezi 1–1,5, byly považovány za

rytmicky velmi podobné nebo zcela identické. Takové hodnocení bylo očekávatelné u položek, které obsahovaly mluvní takt s jeho vlastním nízkofrekvenčním obrazem, avšak u položek obsahujících dva různé takty nebo takt a nízkofrekvenční obraz jiného taktu bylo takovéto hodnocení relativně vzácné – celkem toto kritérium splňovalo 11 položek.

Osm položek z této skupiny se vyznačovalo přítomností rozdílu v sonoritě konsonantů. Značná převaha této vlastnosti v rámci položek s takty hodnocenými jako rytmicky téměř identické je v souladu s obecnými výsledky, které ukázaly, že rozdíl v sonoritě konsonantů v rámci jedné položky je subtilní, či spíše nulový; naopak položky pouze s tímto rozdílem měly tendenci být hodnoceny lépe než ostatní (viz kapitola 4.2.4).

položky	takt 1	takt 2	počet hodnocení	medián	rys	dlouhý V
P35	výšce	<i>místy</i>	40	1	sonorita	ano
P71	nimi	<i>bude</i>	40	1	sonorita	ne
P73	sboru	<i>vjely</i>	40	1,5	sonorita	ne
P76	místy	<i>výšce</i>	40	1	sonorita	ano
P80	byli	<i>bude</i>	40	1	sonorita	ne
P169	výšce	<i>místy</i>	40	1	sonorita	ano
P173	jméno	<i>vládě</i>	80	1	sonorita	ano
P184	místy	<i>výšce</i>	40	1,5	sonorita	ano
P4	se zde	<i>Šopko</i>	40	1,5	znělost	ne
P86	se zde	<i>Šopko</i>	40	1	znělost	ne
P11	jsou to	<i>státy</i>	40	1,5	počet C	ano

Tabulka 26. Položky s extrémně nízkým mediánem hodnocení posluchači. Ve sloupci *položky* jsou čísla analyzovaných položek. Ve sloupci *takt 1* je uveden slovní obsah prvního taktu položky; sloupec *takt 2* obsahuje obsah druhého taktu v položce. Text v tomto sloupci psaný kurzívou značí, že druhý takt v položce zazněl jen jako nízkofrekvenční obraz. V dalších sloupcích je uveden celkový počet hodnocení položky a medián jejího hodnocení. Poslední sloupce obsahují rys, kterým se daná dvojice mluvních taktů v položce od sebe lišila, a přítomnost dlouhého vokálu v obou taktech položky.

Další dvě položky, které byly posluchači považovány za rytmicky ekvivalentní, obsahovaly mluvní takty, jež se lišily rysem znělosti šumových konsonantů. Tento rozdíl se v rámci obecných analýz ukázal jako nepodstatný. Poslední položka se lišila počtem konsonantů ve shluku. Rozdíl v pozici souhláskového shluku v taktech v položce se v rámci rytmicky velmi podobných položek neobjevil vůbec, což nepřímo potvrzuje jeho důležitou roli pro vnímání řečového rytmu.



Podle souhrnných výsledků měly být takty s dlouhým vokálem hodnoceny jako podobnější než takty se všemi samohláskami krátkými (viz kapitola 4.2.3). Z 11 dvojic taktů hodnocených jako extrémně rytmicky podobné jich obsahovalo dlouhý vokál šest. To sice na první pohled nevypadá jako zvlášť silný rys, nicméně celkově bylo v testu pouze 69 položek, které obsahovaly dlouhý vokál, oproti 154 položkám se všemi vokály fonologicky krátkými.

Segmentální obsah – tedy slova, která tvořila takty v jednotlivých položkách – nebyl ve všech 11 položkách unikátní. Rozdílných dvojic taktů bylo pouze sedm; dvojice [vi:ʃtɛ] a [mi:stɪ] se v různých obměnách (pořadí taktů; plné takty, či nízkofrekvenční obraz) v těchto jedenácti položkách objevila celkem čtyřikrát, dvojice [sɛ zdɛ] a [ʃopko] se ve stejném výčtu objevila dvakrát.

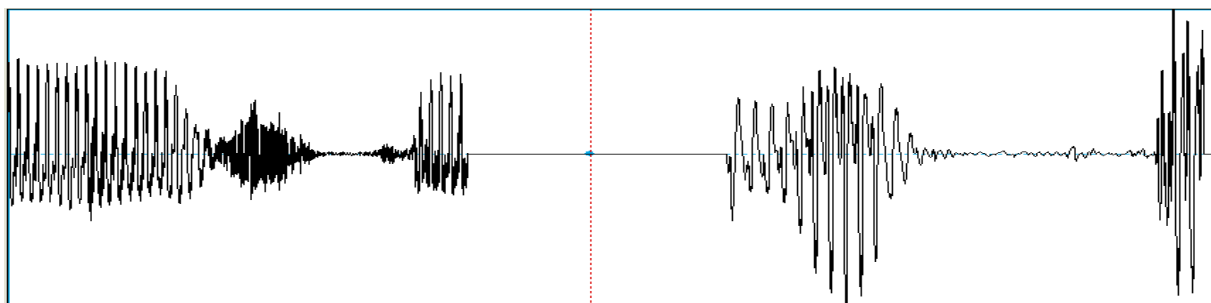
Dvojice [vi:ʃtɛ] a [mi:stɪ] byla hodnocena jako velmi rytmicky si podobná v následujících případech:

pol.	takt 1	takt 2	počet hodn.	medián	modus	trvání 1 (ms)	trvání 2 (ms)	rozdíl trvání (ms)	rozdíl trvání %
P35	výšce	<i>místy</i>	40	1	1	370,02	347,62	22,4	6,05
P76	místy	<i>výšce</i>	40	1	1	347,62	370,02	22,4	6,44
P169	výšce	<i>místy</i>	40	1	1	370,02	347,62	22,4	6,05
P184	místy	<i>výšce</i>	40	1,5	1	347,62	370,02	22,4	6,44

Tabulka 27. Položky hodnocené jako velmi si rytmicky podobné a obsahující dvojici [vi:ʃtɛ] a [mi:stɪ].

Na rytmické podobnosti této dvojice se posluchači shodovali nezávisle na pořadí položek (položky P35 a P169 mají jiné pořadí taktů než položky P76 a P184), přítomnosti nízkofrekvenčního obrazu (v položkách P35, P76 a P169 je druhý takt jako nízkofrekvenční obraz; v položce P184 zazněl plně) i manipulace s intonací (položka 184 má na rozdíl od zbylých tří uměle vyhlazenou intonační linii). Rytmická podobnost těchto konkrétních taktů byla natolik nápadná, že byly prozkoumány i ostatní položky s touž dvojicí v testu, které nedosahovaly extrémního hodnocení posluchači, a bylo zjištěno, že ani u ostatních takto „sesterských“ položek nebyl nikdy medián hodnocení vyšší než 2 a modus byl dokonce vždy roven jedné. Je tedy záhodno nastínit možné interpretace tohoto jevu. Jako první vodítko může posloužit rozdíl v trvání taktů v položce, který tvořil pouze 22,4 ms, procentuálně vyjádřeno pak mezi 6–6,5 % trvání prvního taktu z dvojice. Další příčinou hodnocení této dvojice jako rytmicky podobné může také být hláskové složení porovnávaných taktů. V kapitole 4.2.4 byl vliv kontrastu

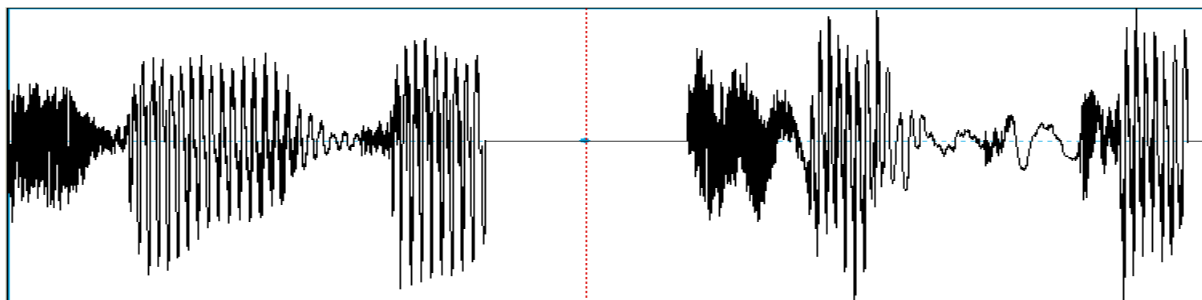
sonority znělých souhlásek na vnímání rytmické podobnosti označen jako subtilní. Položky s rozdílem v sonoritě konsonantů byly totiž hodnoceny spíše rytmicky podobněji než ostatní, které mnohdy obsahovaly percepčně výraznější rozdíly (např. pozice konsonantického shluku v taktu). Takty, které se vyznačovaly pouze fonologickým rozdílem v sonoritě znělých konsonantů, se ve skupině položek s nízkým mediánem hodnocení vyskytly osmkrát z celkových 11 položek (zatímco například položky s rozdílem znělosti konsonantů jen ve dvou takových případech). Kapitola 4.2.3 také poukázala na vliv přítomnosti dlouhého vokálu ve dvojici taktů na hodnocení rytmické podobnosti – položky s první slabikou obsahující dlouhou samohlásku byly hodnoceny jako podobnější si navzájem než položky, v nichž se vyskytovaly pouze krátké vokály. Oba takty ve dvojici [vi:ʃtɛ] a [mi:stɪ] mají také velmi podobnou strukturu co do časových proporcí (průběh základního tónu na začátcích taktů trvá 140–150 ms, následuje 160–180 ms nepřítomnosti  $f_0$  a poté cca 50 ms opět se základním tónem). Energie bývá relativně silná i u neznělých ostrých sykavek, do vnímání rytmu však pravděpodobně tyto hlásky příliš nezasahují: ačkoli se do nízkofrekvenčního obrazu se totiž neznělé sibilanty promítly pouze jako ticho (obr. 4), byly tyto dvojice hodnoceny jako rytmicky téměř ekvivalentní.



Obr. 4. Oscilogramy taktu [mi:stɪ] a nízkofrekvenčního obrazu taktu [vi:ʃtɛ]. Na oscilogramech tohoto i dalších nízkofrekvenčních obrazů je patrný větší rozsah amplitudy než u jejich plně proslovených protějšků: mluvním taktům s odstraněnými vyššími frekvencemi byla totiž kvůli percepční srovnatelnosti obou taktů ve dvojici hlasitosti zvýšena intenzita o 5 dB.

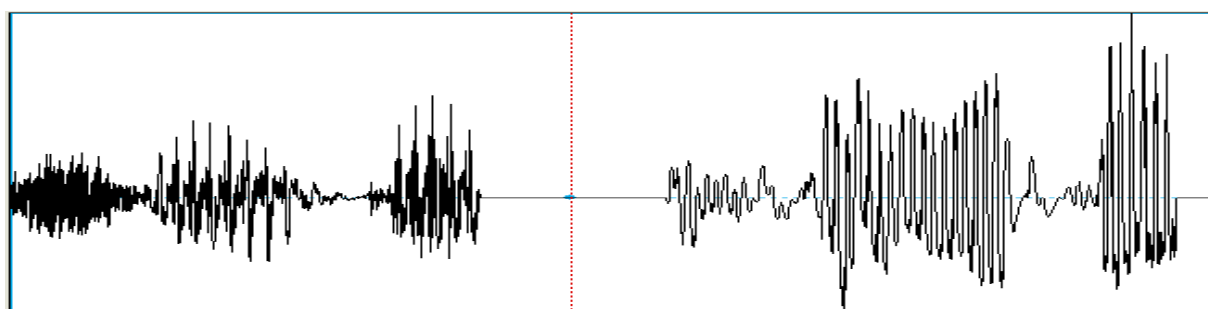
Další položky se stejnou dvojicí taktů byly P4 a P86 [ɛ zɛ]–[ʃopko]. Tyto položky byly dvě a v obou případech zachovávaly uvedené pořadí taktů, avšak jedna položka obsahovala druhý takt ve formě nízkofrekvenčního obrazu a ve druhé se vyskytovaly dva plné takty. Jako rytmicky podobnější byla poněkud překvapivě hodnocena položka se dvěma plnými takty (medián hodnocení = 1 na rozdíl od 1,5 u položky s nízkofrekvenčním obrazem), ačkoli obecně byly jako podobnější hodnoceny položky

s nízkofrekvenčními obrazy (viz kapitola 4.1.2). Oba takty nicméně měly téměř shodné trvání – rozdíl tvořilo pouhých 12,7 ms, což tvořilo 4,2 % z trvání prvního taktu v položce. Také konsonanticko-vokální struktura byla u obou taktů podobná: CVCCV. Rozdíl ve znělosti souhláskového shluku uprostřed taktu ([zd] versus [pk]) byl posluchači vyhodnocen jako nepříliš podstatný pro vnímání rytmu, jak se ostatně ukázalo již v souhrnné analýze v kapitole 4.2.5. Oba takty tedy mohly být posluchači vnímány jako dva silné pulzy oddělené zdánlivým tichem (obr. 5).



Obr. 5. Oscilogramy taktů [se zde] a [fopko].

Z položek, které obsahovaly různé takty a zároveň byly hodnoceny jako rytmicky si velmi podobné, pouze dvě tvořily dva plně vyslovené mluvní takty: u ostatních devíti se tedy vždy vyskytoval nízkofrekvenční obraz slova, který obvykle spíše nivelizoval rozdíly mezi takty. Jednalo se o položky P86 a P184. Obě položky obsahovaly již výše zmíněné dvojice taktů [se zde]–[fopko] a [vi:řtse]–[mi:stɪ].

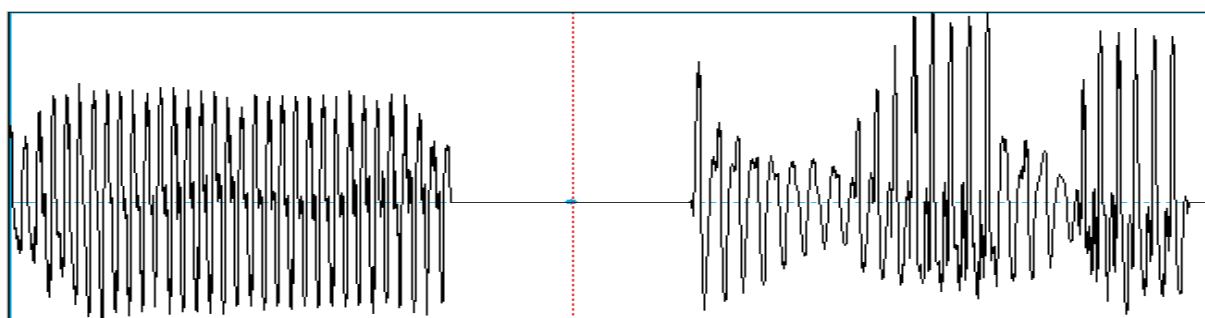


Obr. 6. Oscilogramy taktu [sôu to] a nízkofrekvenčního obrazu taktu [sta:ti].

Položka P11, rovněž hodnocená jako obsahující rytmicky podobné takty, obsahovala dvojici [sôu to]–[sta:ti] s druhým taktem coby nízkofrekvenčním obrazem. Tyto dva takty vykazují co do zkoumaných vlastností pouze jediný rozdíl, a to počet konsonantů v první souhláskové skupině. Jelikož jsou všechny hlásky na začátcích prvních slabik taktů neznělé, po odstranění vysokých frekvencí je rozdíl mezi takty percepčně obtížně zachytitelný. Rozdíly v časových proporcích první slabiky u obou taktů jsou také

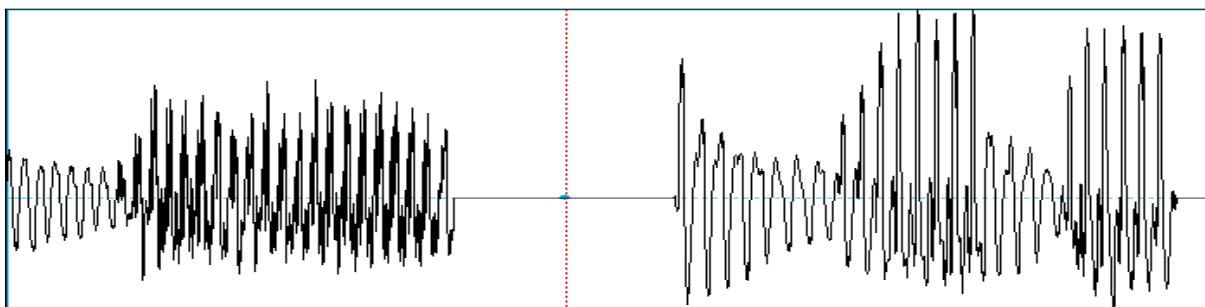
zanedbatelné: jak [s] u [sôu to], tak [st] u [sta:ti] trvá cca 70–80 ms; rozdíl v trvání vokálu je pouze cca 20 ms, přičemž celkový rozdíl v trvání taktů je 26,5 ms, což tvoří 9,2 % trvání prvního taktu. Také se zde potvrzuje zjištění z kapitoly 4.2.3, v níž se uvádí, že položky s první slabikou obsahující dlouhou samohlásku byly hodnoceny podobněji než položky se samými krátkými vokály.

Další dvojici taktů hodnocenou jako rytmicky podobnou, byla položka P71 obsahující takt [ɲimɪ] a nízkofrekvenční obraz taktu [bude] (obr. 7). Zde bylo hodnocení rytmické podobnosti očekávatelné vzhledem k subtilnosti rozdílu sonorních a znělých šumových souhlásek, na což bylo poukázáno v kapitole 4.2.4. Při pohledu na oscilogram jsou však patrné rozdíly mezi oběma takty: zatímco u prvního taktu [ɲimɪ] slabičné vrcholy nejsou prakticky nijak oddělitelné, u nízkofrekvenčního obrazu taktu [bude] jsou jasně patrné. V percepci jsou ale slabiky u obou taktů odlišitelné relativně zřetelně. Co do trvání jsou si oba takty podobné, i když ve srovnání s předchozími položkami je rozdíl vyšší: takty v položce se liší o 29,1 ms, což tvoří 13,2 % trvání prvního taktu.



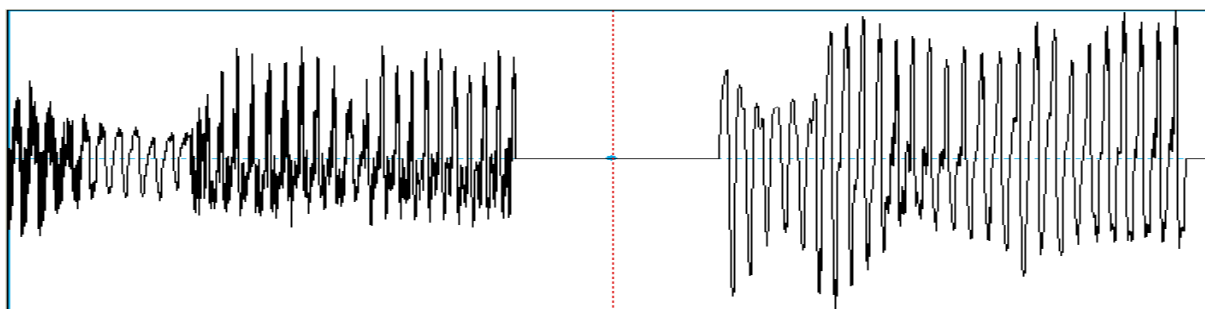
Obr. 7. Oscilogramy taktu [ɲimɪ] a nízkofrekvenčního obrazu taktu [bude].

Velmi podobný případ jako dvojice [ɲimɪ]–[bude] je i položka P80 obsahující takt [bɪlɪ] a nízkofrekvenční obraz taktu [bude] (obr. 8). Také zde se projevuje nízký význam rozdílu v rysu sonority u sonorních a znělých šumových hlásek. Slabičné vrcholy jsou v oscilogramu prvního taktu nepříliš zřetelně oddělené, percepčně jsou však odlišitelné dostatečně. Časové proporce jsou téměř stejné jako u dvojice [ɲimɪ]–[bude]: rozdíl v celkovém trvání taktů je 28 ms nebo 12,7 % z trvání prvního taktu.



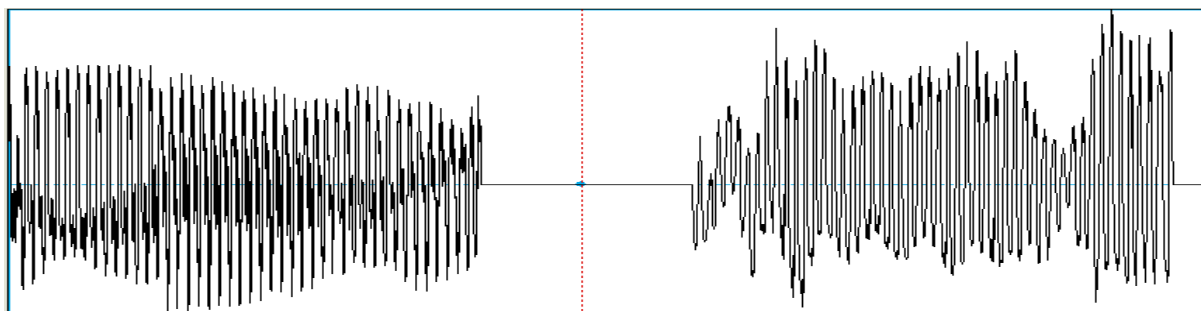
Obr. 8. Oscilogramy taktu [bɪlɪ] a nízkofrekvenčního obrazu taktu [bude].

Mluvní takt [zboru] a nízkofrekvenční obraz taktu [vjelɪ] (obr. 9) v rámci položky P73 pravděpodobně vděčí za hodnocení rytmické podobnosti taktů subtilnímu rozdílu mezi sonorními a znělými nesonorními konsonanty. Oba takty mají strukturu CCVCV; liší se pouze sonoritou konsonantického shluku v préture první slabiky. Rozdíl v trvání taktů je zanedbatelný (13,4 ms, 5,3 % z trvání prvního taktu).



Obr. 9. Oscilogramy taktu [zboru] a nízkofrekvenčního obrazu taktu [vjelɪ].

Položka P173 obsahující takt [jmɛ: no] a nízkofrekvenční obraz taktu [vɫa: jɛ] opět potvrzuje předchozí zjištění obecných analýz i výše uvedené položkové analýzy: Položky s rozdílem v sonoritě konsonantů byly v rámci tohoto testu hodnoceny rytmicky podobněji než ostatní, které často obsahovaly percepčně výraznější rozdíly, a dvojice taktů s první slabikou obsahující dlouhou samohlásku byly hodnoceny jako podobnější než položky, v nichž se objevily pouze krátké vokály. Tato položka patřila k těm, u nichž byla manuálně vyhlazena intonační linie, nicméně vliv vyhlazení melodie prokázán nebyl (viz kapitola 4.3.1).



Obr. 10. Oscilogramy taktu [jme: no] a nízkofrekvenčního obrazu taktu [vla: ɛ].

#### 4.4.3.2 *Položky s vysokým mediánem hodnocení*

Vysoký medián hodnocení (4,5–5) položky v percepčním testu indikoval, že respondenti zvuky v dané položce hodnotili jako rytmicky výrazně odlišné. Takových položek bylo celkem 10. U položek, které obsahovaly plně proslovený takt a jeho vlastní nízkofrekvenční obraz, byly za vysokou hodnotu u mediánu hodnocení položky považovány hodnoty vyšší než 1,5. Těchto položek se ve vyhodnocení vyskytovalo celkem sedm.

##### 4.4.3.2.1 *Položky obsahující rozdílné takty*

Ačkoli se v rámci položek hodnocených jako rytmicky výrazně odlišných nevyskytovala dvojice taktů, která by se v různých obměnách opakovala, jako tomu bylo u položek s nízkým mediánem hodnocení, bylo možno i v této skupině vysledovat jisté pravidelnosti. Všechny položky obsahovaly nízkofrekvenční obraz taktu – to znamená, že mezi nimi nebyla ani jedna, která by byla tvořena dvěma plně proslovenými takty. Nejvíce položek v této skupině se vyznačovalo rozdílem ve vlastnosti, která se ukázala jako stěžejní pro vnímání rytmické podobnosti – v pozici konsonantického shluku v taktu (tabulka 28).

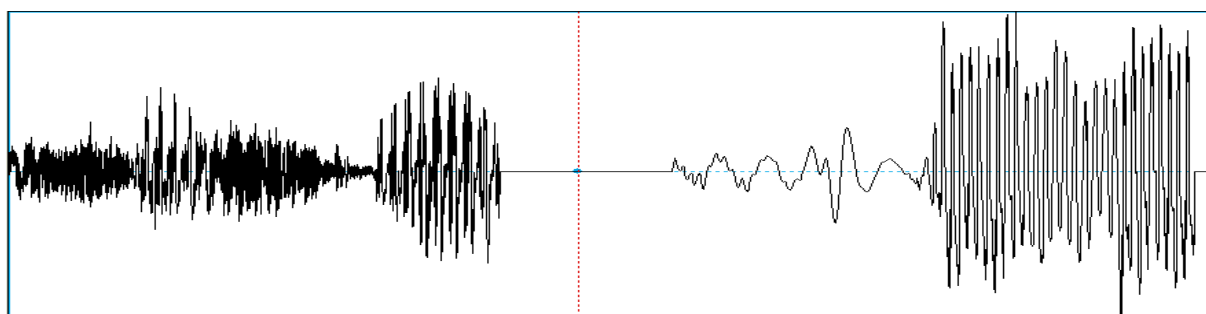
Dalším rysem, který figuroval u tří položek s rozdílnými takty a s vysokým mediánem hodnocení, byl poněkud překvapivě rozdíl v sonoritě, který by naopak významným pro vnímání rytmu být neměl (viz kapitola 4.2.4). Bližší zkoumání ale u těchto položek odhalilo silnější odlišnosti, které pravděpodobně extrémní hodnocení zapříčinilo (např. vložený slabičný prvek). Dvě položky z této skupiny dále obsahovaly rozdíl v počtu konsonantů ve shluku a v taktech jedné položky se objevil i kontrast znělosti.

položky	takt 1	takt 2	počet hodnocení	medián	rys	dlouhý V
P23	sešlo	škody	40	4,5	pozice	ne
P63	době	sboru	40	4,5	pozice	ne
P66	jedna	vjely	40	4,5	pozice	ne
P68	jedna	sboru	40	5	pozice	ne
P30	bude	nimi	40	5	sonorita	ne
P34	v noci	Lhase	40	4,5	sonorita	ne
P82	mezi	nimi	40	5	sonorita	ne
P10	bude	v Gaze	40	5	počet C	ne
P56	země	nimi	40	5	počet C	ne
P2	státy	svého	40	4,5	znělost	ano

Tabulka 28. Položky s extrémně vysokým mediánem hodnocení posluchači. Ve sloupci *položky* jsou čísla analyzovaných položek. Ve sloupci *takt 1* je uveden slovní obsah prvního taktu položky; sloupec *takt 2* obsahuje obsah druhého taktu v položce. Text v tomto sloupci psaný kurzívou značí, že druhý takt v položce zazněl jen jako nízkofrekvenční obraz. V dalších sloupcích je uveden celkový počet hodnocení položky a medián jejího hodnocení. Poslední sloupce obsahují rys, kterým se daná dvojice mluvních taktů v položce od sebe lišila, a přítomnost dlouhého vokálu v obou taktech položky.

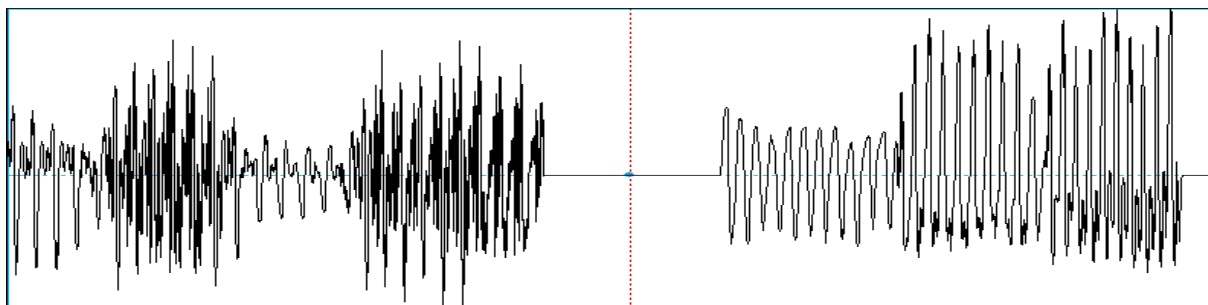
Položky s takty obsahujícími dlouhou samohlásku byly hodnoceny jako rytmicky si podobnější (viz kapitola 4.2.3), čemuž odpovídá i fakt, že mezi 10 dvojicemi taktů, které posluchači označili jako rytmicky téměř či zcela odlišné, takty s dlouhou samohláskou figurovaly jen v jediném případě.

Položka P23 obsahovala dvojici [seflo]–[fkodɪ] (obr. 11), která se fonotakticky odlišovala pozicí souhláskového shluku. Rozdíl v trvání mezi jedním taktem a nízkofrekvenční variantou taktu druhého nebyl nijak vysoký (24 ms; 7,3 % trvání prvního taktu), položka však celkově dosáhla mediánu hodnocení 4,5. Jednou z příčin může být výraznější rozdíl v trvání úseku mezi intenzitními vrcholy vlivem přítomnosti, či nepřítomnosti konsonantické skupiny: zatímco v prvním taktu jsou vrcholy od sebe vzdáleny cca 110 ms, v obrazu druhého taktu je časová vzdálenost mezi těmito vrcholy jen cca 60 ms.



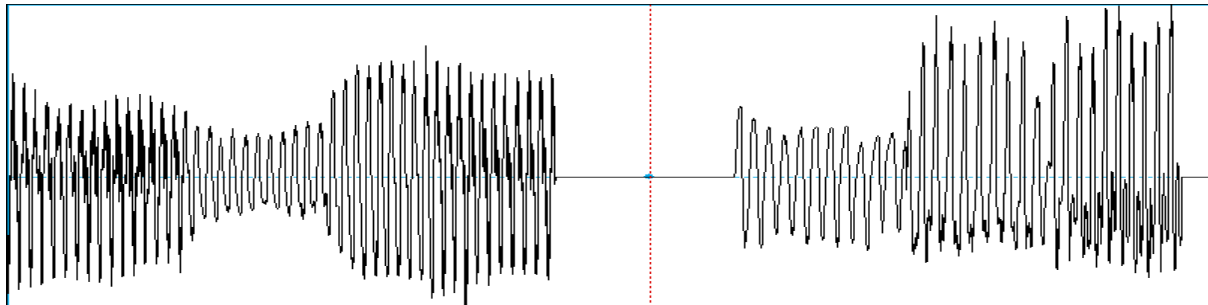
Obr. 11. Oscilogramy taktu [seflo] a nízkofrekvenčního obrazu taktu [fkodɪ].

Podobný rozdíl v trvání části mluvního taktu mezi dvěma intenzitními vrcholy je možné pozorovat i u položky P63, obsahující dvojici taktů [dobje]–[zboru] (obr. 12). Zatímco vzdálenost intenzitních vrcholů v taktu [dobje] činí cca 80 ms, u obrazu taktu [zboru] tvoří úsek mezi vrcholy pouze cca 30 ms.



Obr. 12. Oscilogramy taktu [dobje] a nízkofrekvenčního obrazu taktu [zboru].

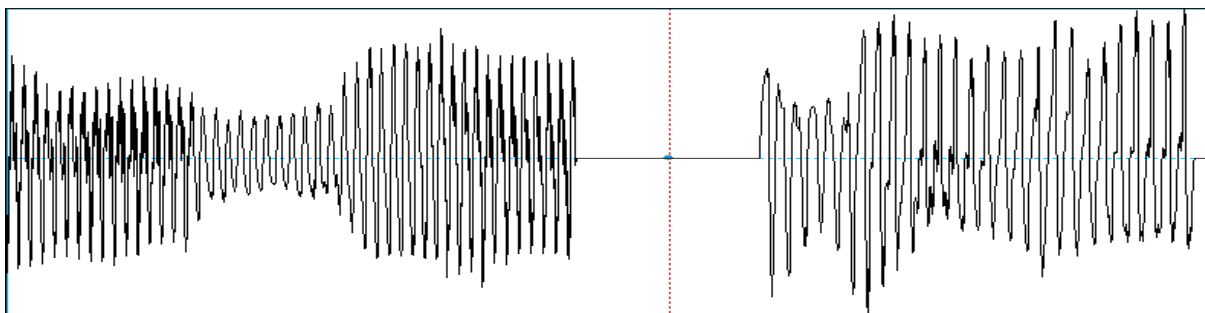
Položka P68 s dvojicí taktů [jedna]–[zboru] (obr. 13) vykazuje také výrazný rozdíl v trvání úseku mezi dvěma intenzitními vrcholy: u prvního taktu tento úsek tvoří bezmála 100 ms, zatímco u druhého cca 30 ms. Také se zde na rozdíl od předchozích dvou položek vyskytuje větší rozdíl v trvání jednotlivých zvuků – 51,3 ms (16,9 % trvání prvního taktu v položce).



Obr. 13. Oscilogramy taktu [jedna] a nízkofrekvenčního obrazu taktu [zboru].

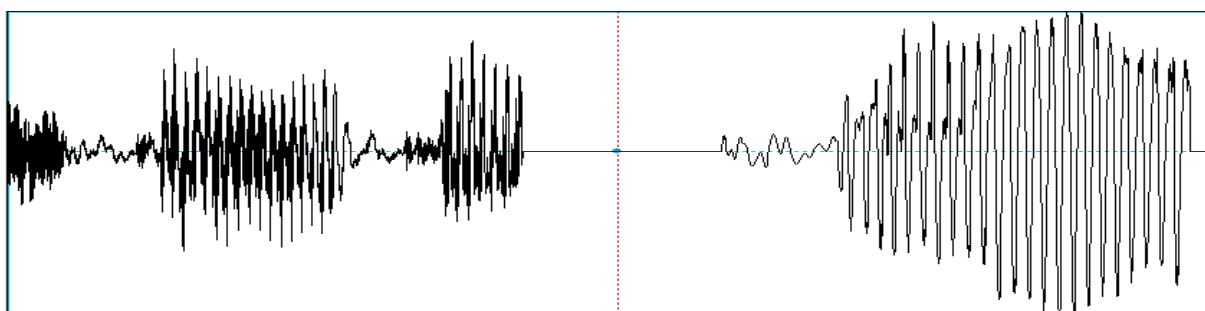
Položka P66 s dvojicí [jedna]–[vjeɫɪ] (obr. 14) se může svou konsonanticko-vokalickou strukturou jevit jako případ ekvivalentní třem výše uvedeným, a jak je také patrné z oscilogramů na obr. 14, u nízkofrekvenčního obrazu taktu [vjeɫɪ] není pokles mezi jednotlivými intenzitními vrcholy zcela zřetelný. Význam však může nést i značný rozdíl v trvání obou taktů: 64,7 ms, 21,3 % z trvání prvního taktu v položce.





Obr. 14. Oscilogramy taktu [jedna] a nízkofrekvenčního obrazu taktu [vjelɪ].

Silná sonorní složka je pravděpodobně příčinou vysokého mediánu hodnocení položky P2 obsahující dvojici [sta:ti]–[svɛ:ɦo]. Ačkoli kontrast znělosti nesehrál při celkovém hodnocení rytmické podobnosti významnou roli (viz kapitola 4.2.5.), tento případ může tvořit výjimku. V prvním taktu jsou totiž intenzitní vrcholy odděleny neznělými hláskami, zatímco u druhého je obtížné jednotlivé vrcholy zřetelně odlišit v oscilogramu i percepci (obr. 15). Rozdíl trvání obou zvuků je přitom zanedbatelný (25,5 ms; 8,1 % trvání prvního taktu v položce).

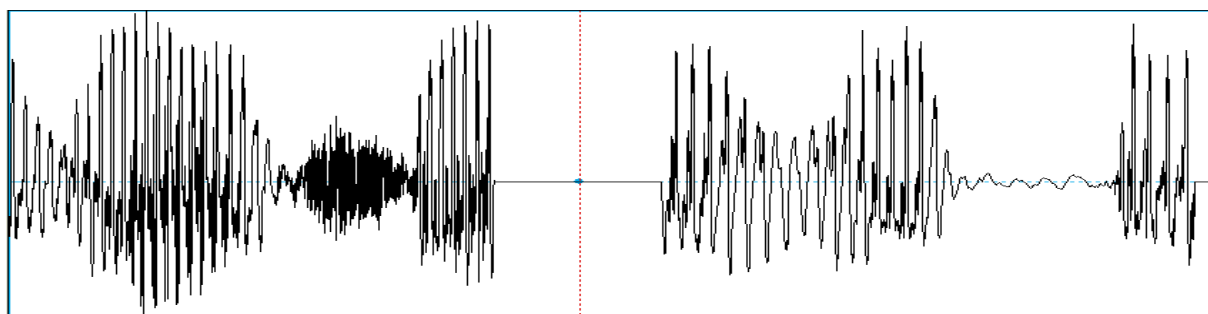


Obr. 15. Oscilogramy taktu [sta:ti] a nízkofrekvenčního obrazu taktu [svɛ:ɦo].

Takty v položce P34, obsahující dvojici [v noʃɪ]–[lɦasɛ], vykazovaly podobný (tedy malý) rozdíl v trvání jako položka předchozí. Fonotaktický rozdíl mezi oběma takty byl v záměně pořadí sonorní a znělé hlásky v iniciálním konsonantickém shluku v obou taktech [vn]–[lɦ]. Bližší zkoumání odhalilo také vnímatelnou odlišnost v intonačním průběhu zvuků, která se projevila při měření rozdílů výšky tónu mezi první a druhou slabikou u každého mluvního taktu. Rozdíly ve frekvenci jednotlivých slabik v souvislosti s vnímanou výškou byly měřeny v polovině jádra slabiky, protože právě zde se vjem výšky slabiky vztahuje k frekvenci (Beckman, 1986; Hermes, 2006; Volín, 2009). Zatímco u taktu [v noʃɪ] byla intonační linie monotónní (rozdíl v melodii mezi středem jádra první a druhé slabiky v prvním taktu činil pouze 0,22 půltónu), druhý takt vykazoval melodickou kadenci stoupavě-klesavou (střed jádra první slabiky byl

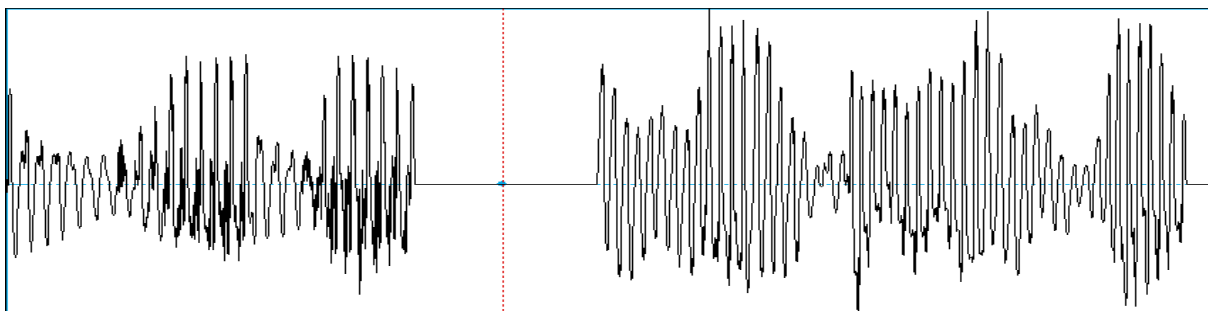
o 4,04 pŭltónu výš než jádro slabiky druhé). Pravděpodobně rozhodujícím faktorem pro ohodnocení taktů jako rytmicky odlišných však byla přítomnost krátkého středního středového vokálu uprostřed souhláskové skupiny [lɦ] ve druhém taktu. V mluvních taktech obsahujících uspořádání hlásek sonora-obstruent-vokál v rámci jedné slabiky není tento jev ničím výjimečný. Takto vytvořený vedlejší sonorní vrchol bývá nazýván pobočnou slabikou (např. Palková, 1994). Pobočná slabika porušuje kvaziuniverzální princip vzrůstající sonority ve slabice, podle kterého sonorita vzrůstá od začátku k vrcholu slabiky a opět klesá od jádra k jejímu konci (např. Hála, 1956; Clements, 1990; Goldsmith, 2011).

Konsonantický shluk [lɦ] byl tedy v uvedeném případě reálně vysloven jako [ləɦ], což výrazně vyvstalo po odstranění segmentálního obsahu taktu. Ten měl tak ve svém nízkofrekvenčním obrazu tři intenzitní vrcholy (obr. 16).



Obr. 16. Oscilogramy taktu [v noɦsɪ] a nízkofrekvenčního obrazu taktu [lɦase].

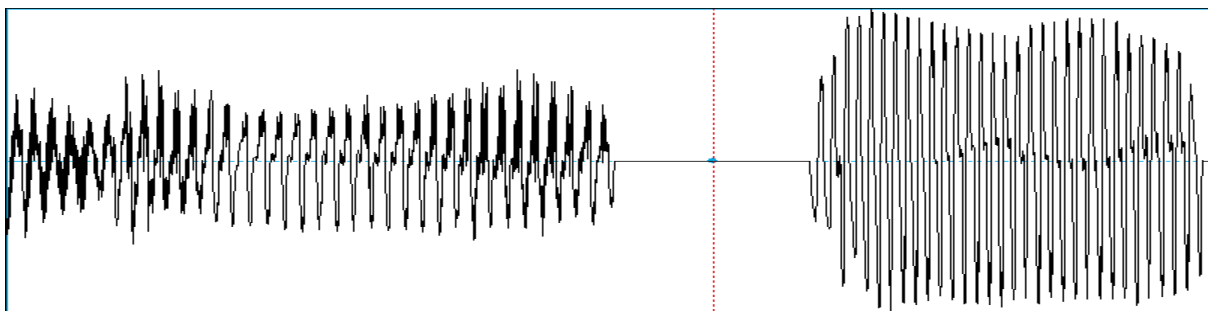
Epentetická střední středová samohláska alias schwa se v iniciálním souhláskovém shluku objevila také v položce P10 v nízkofrekvenčním obrazu taktu [v gaze], který tvořil dvojici s taktem [bude] (obr. 17). Kvůli přítomnosti této samohlásky se formálně dvouslabičný mluvní takt změnil v tříslabičný, v němž první slabika obrazu taktu [v gaze] byla slabika [və] s vlastním intenzitním vrcholem. Vkládání epentetického schwa je v češtině jev poměrně častý, objevuje se více v pečlivější řeči a souvisí se způsobem i místem tvoření souhlásek: podle Hály (1956) je vokalický prvek tím výraznější, čím složitější je přenastavení mluvidel mezi dvěma konsonanty. V tomto případě se také projevil velmi výrazný rozdíl v trvání obou taktů, který činil 113,1 ms, tedy 45,4 % trvání prvního taktu. Kromě přidané samohlásky totiž druhý takt na rozdíl od prvního obsahoval souhláskový shluk.



Obr. 17. Oscilogramy taktu [bude] a nízkofrekvenčního obrazu taktu [v gaze].

Pozoruhodným faktem bylo, že položka obsahující tytéž takty ve stejném pořadí, avšak plně proslovené, vykazovala medián hodnocení 2, tedy rytmicky relativně podobné. Je tedy možné, že střední středová samohláska v konsonantickém shluku není percepčně tolik významná, pokud je přítomen segmentální obsah taktu. Že je přítomnost epentetického [ə] v některých případech nevyhnutelná a nemusí v projevu působit rušivě, pokud nedává vzniknout dojmu přidané slabiky, konstatovali už Skarnitzl a Machač (2012). V uvedeném případě dojem nadbytečné slabiky vzniká až po odstranění vyšších frekvencí.

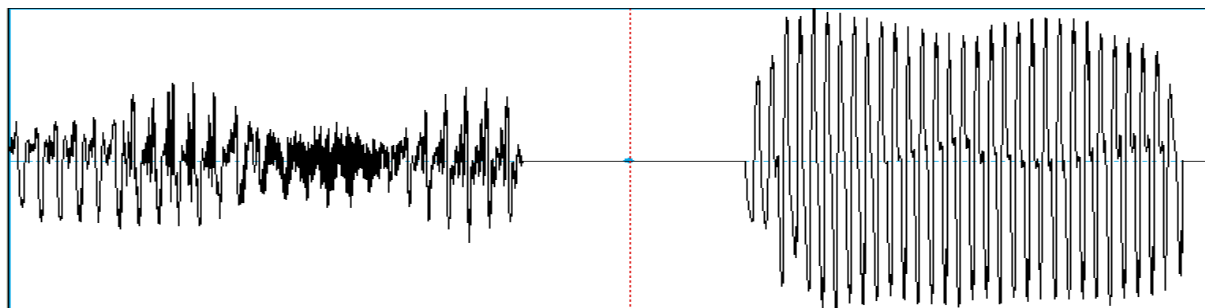
Odlišnost v přítomnosti konsonantické skupiny vykazovala dvojice taktů v položce P56 [zɛmpɛ]–[ɲimɪ], tentokrát se však rozdíl v počtu souhlásek týkal mediální pozice v mluvním taktu. Oba takty mají podobné rozložení intenzity bez zřetelných předělů v podobě poklesu intenzity (obr. 18). Čím se ale liší, je trvání celého taktu (rozdíl 112,9 ms, tedy 33,9 % trvání prvního taktu) a trvání sonorní části taktu (u taktu [zɛmpɛ] trvá sonorní část cca 275 ms, u taktu [ɲimɪ] pouze 210 ms).



Obr. 18. Oscilogramy taktu [zɛmpɛ] a nízkofrekvenčního obrazu taktu [ɲimɪ].

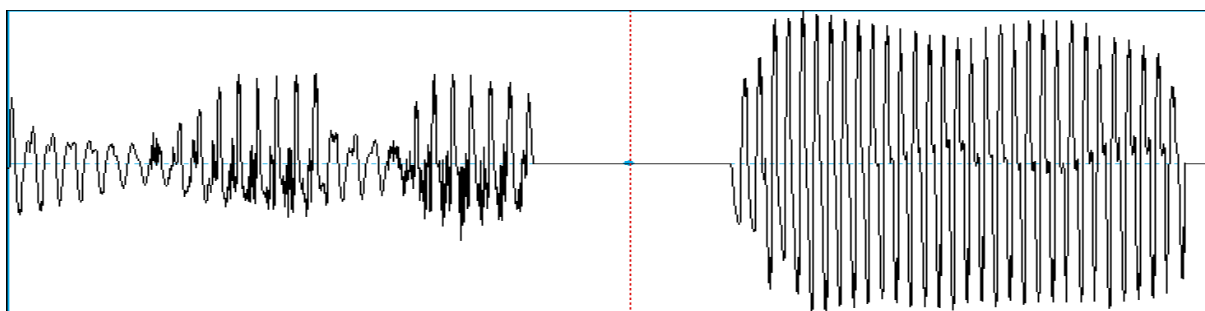
V percepčním testu byly také jako velmi rytmicky odlišné ohodnoceny položky P30 a P82, které obsahovaly nízkofrekvenční obraz taktu [ɲimɪ], stejně jako položka P56. Tento takt se v položce P56 objevil vedle taktu [zɛmpɛ], který se lišil počtem konsonantů v mediálním shluku; v položkách P30 a P82 se však odlišoval pouze rozdílem v sonoritě

souhlásek. V případě dvojice [mezi]–[nimi] v položce P82 šlo o kontrast [z]–[m], kde byla navíc frikativa částečně desonorizována (ztráta znělosti u hlásky [z] je v češtině relativně častá; např. v Macháčově studii (2008) se objevila ve 33,3 % případů ze všech realizací hlásky; o desonorizaci také viz Skarnitzl, 2011).



Obr. 19. Oscilogramy taktu [mezi] a nízkofrekvenčního obrazu taktu [nimi].

Ve dvojici [bude]–[nimi] v položce P30 šlo pak o rozdíl v sonoritě obou konsonantů v taktech: [b]–[ɲ] a [d]–[m]. Rozdíl v trvání taktů v obou dvojicích je relativně malý – ani v jednom z případů nepřekračuje 32 ms a 12,5 % trvání prvního taktu ve dvojici.



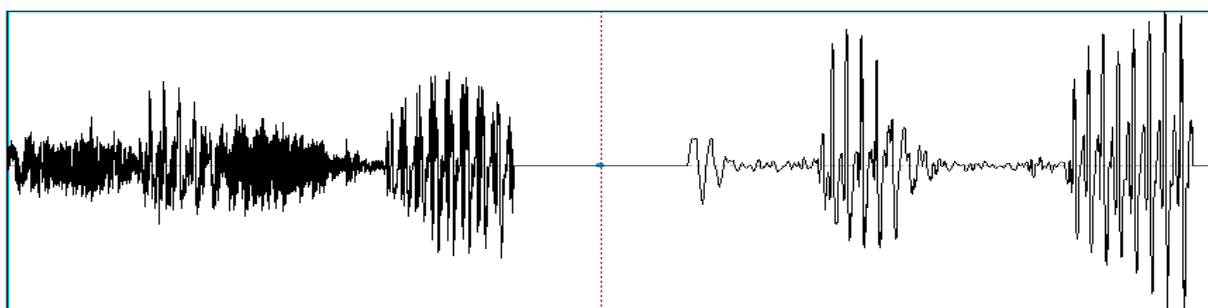
Obr. 20. Oscilogramy taktu [bude] a nízkofrekvenčního obrazu taktu [nimi].

První takty v obou dvojicích mají i další společnou vlastnost – jejich intenzitní vrcholy jsou v percepci i v oscilogramu (obr. 19, 20) zřetelně oddělené, zatímco v taktu [nimi] dochází jen k drobnému poklesu intenzity, takže je dokonce možné, že byl obraz taktu [nimi] vnímán jako jedna slabika s dlouhým vokálem. Tuto hypotézu podporuje i fakt, že takt [nimi] byl s 219,9 ms po taktu [jako] druhým nejkratším taktem co do celkového trvání. S přihlédnutím k poznatkům Weingartové (2015) o trvání hlásek v češtině je možné konstatovat, že trvání celého tohoto taktu by mohlo odpovídat i jediné slabice o třech, nebo dokonce pouze dvou segmentech.

#### 4.4.3.2.2 Položky obsahující stejné takty

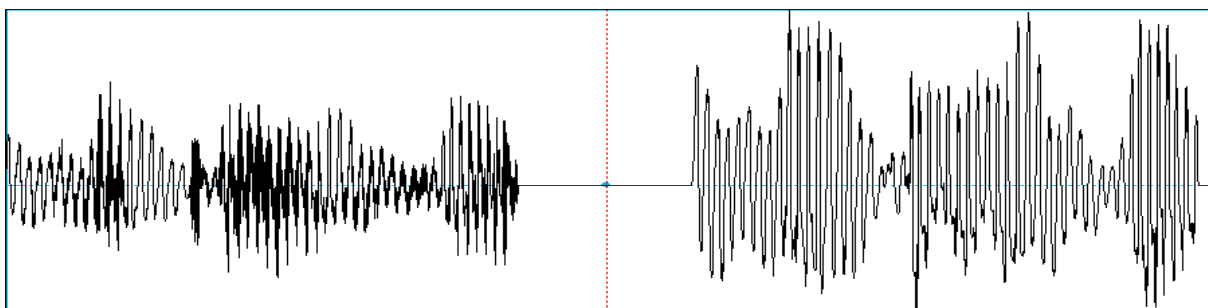
Jednou z hypotéz percepčního testu byl předpoklad, že respondenti budou hodnotit jako rytmicky identické nebo velmi podobné položky s taktem a jeho vlastním nízkofrekvenčním obrazem. Tato hypotéza se potvrdila v analýze popsané v kapitole 4.1.1. Sedm z těchto položek však vykazovalo medián 2, dvojice v nich obsažené tedy byly hodnoceny pouze jako „spíše rytmicky podobné“. Jelikož šlo o identické takty, nemohl zde hrát roli rozdíl v trvání ani v rozdílném průběhu intonační kadence; příčinu jejich hodnocení coby odlišnějších dvojic taktů je tedy nutné hledat jinde.

Položka N16 obsahovala takt [seflo] s jeho nízkofrekvenčním obrazem. Z oscilogramu na obr. 21 je možné vidět, že první intenzitní vrchol v plné realizaci taktu ustupuje silné energii neznělých sibilant, která je však po odstranění vyšších frekvencí zcela potlačena; naopak u obrazu amplituda samohlásky vystupuje do popředí.



Obr. 21. Oscilogramy taktu [seflo] s jeho nízkofrekvenčním obrazem.

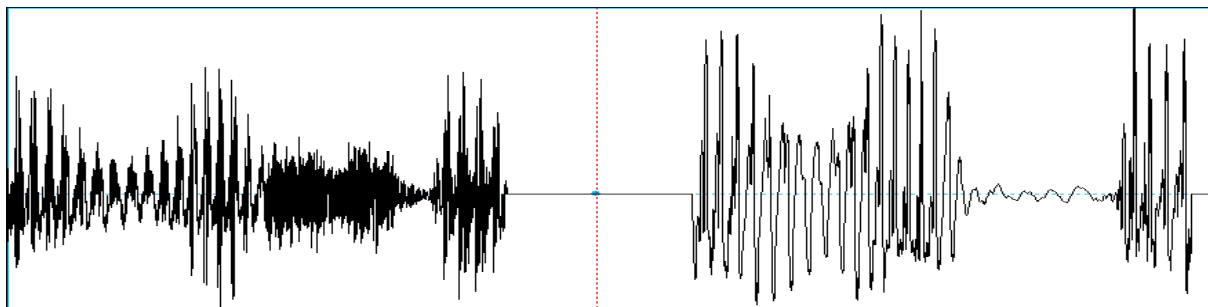
V položce N24 s taktém [v gaze] se pravděpodobně opět projevil vliv epentetické střední středové samohlásky v iniciálním konsonantickém shluku. Vnímání nízkofrekvenčního obrazu jako mírně odlišného od plné výslovnosti mohlo být způsobeno právě absencí segmentálního obsahu, jak už bylo poukázáno v kapitole 4.4.3.2.1.



Obr. 22. Oscilogramy taktu [v gaze] s jeho nízkofrekvenčním obrazem.

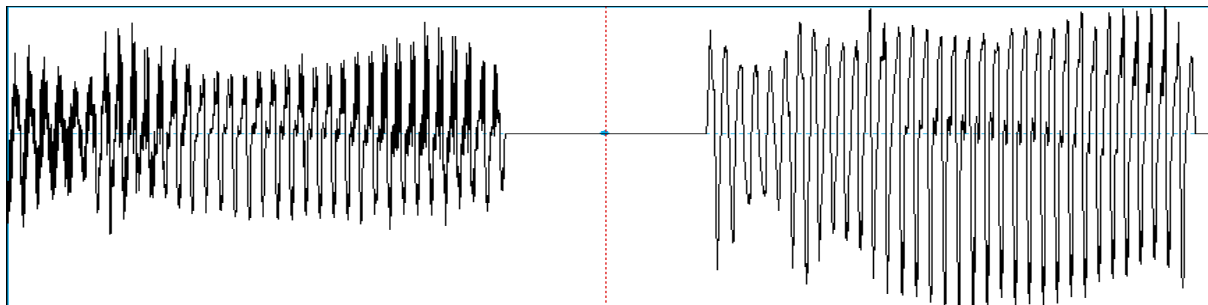
Podobně jako u položky N24 i u taktu [lɦase] (obr. 23) i v položce N30 vyvstala po odstranění vyšších frekvencí nová slabika taktu vzniklá epentetickou střední středovou

samohláskou za prvním konsonantem ve skupině souhlásek. Tuto slabiku pravděpodobně posluchači nevnímali, dokud slyšeli hláskový obsah, a tedy chápali i význam slova.



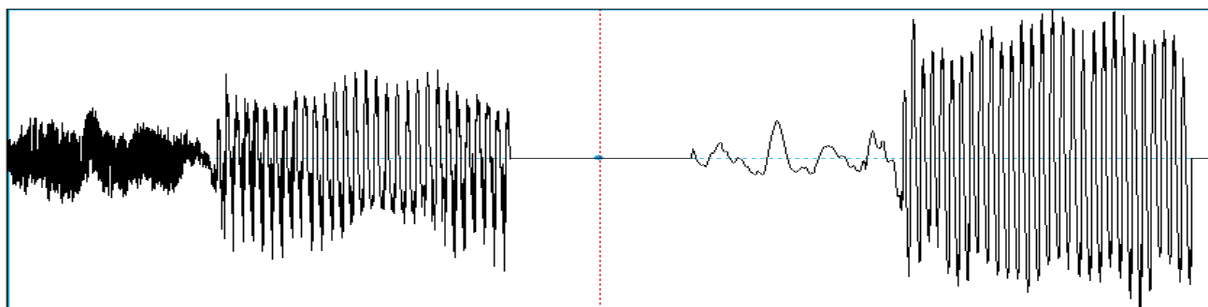
Obr. 23. Oscilogramy taktu [lɦasɛ] s jeho nízkofrekvenčním obrazem.

Položka N26 s dvojicí tvořenou mluvním taktém [zɛmpɛ] a jeho nízkofrekvenčním obrazem (obr. 24) byla hodnocena jako rytmicky rozdílnější zřejmě vlivem sonorní části taktu, která se po odstranění vyšších frekvencí percepčně slila v jediný intenzitní vrchol, typický pro jedinou slabiku. Ačkoli celkové trvání taktu (332,7 ms) spíše ukazovalo na slabiky dvě, respondenti absenci zřetelnějšího předělu mezi intenzitními vrcholy zohlednili v hodnocení dvojice jako rytmicky ne zcela identické.



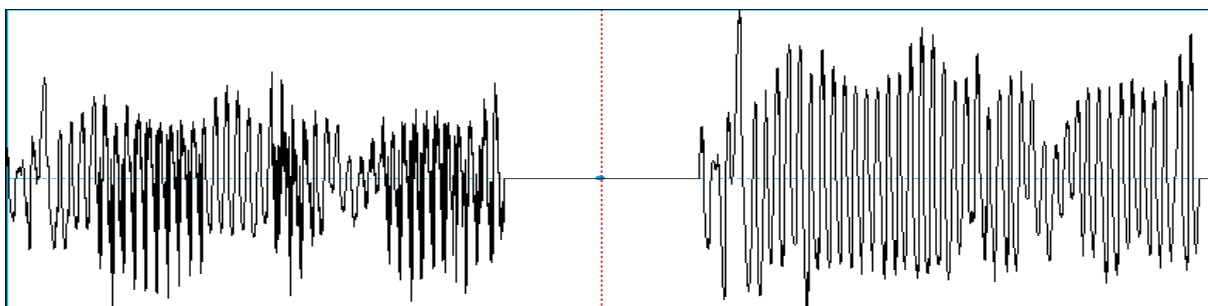
Obr. 24. Oscilogramy taktu [zɛmpɛ] s jeho nízkofrekvenčním obrazem.

Při hodnocení položky N32 s taktém [sɪlɛ] (obr. 25) pravděpodobně sehrál roli též faktor jako v předchozím případě. Sonorní část taktu po odstranění vyšších frekvencí přišla o významnější pokles intenzity, který by oba vrcholy odděloval. Roli však může hrát i fakt, že z trvání celého taktu 303,3 ms tvoří cca 120 ms neznělá sibilanta, která v nízkofrekvenčním obrazu slova téměř zanikla.



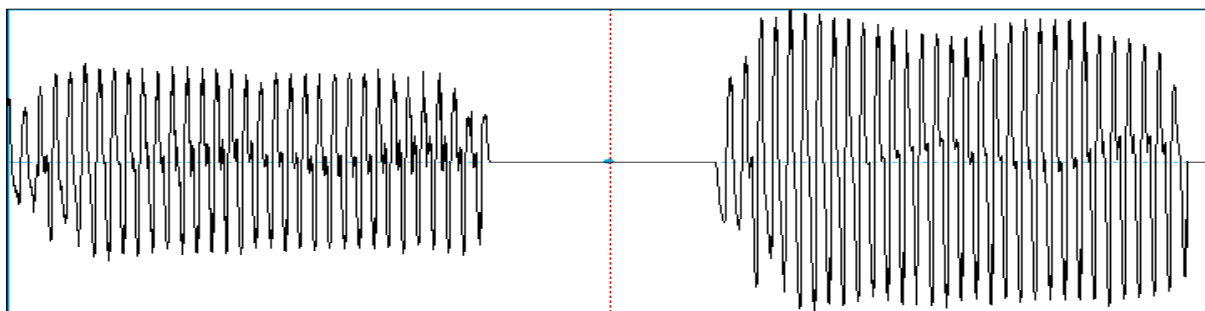
Obr. 25. Oscilogramy taktu [silɲɛ] s jeho nízkofrekvenčním obrazem.

Takt [na dva] a jeho nízkofrekvenční obraz v položce N38 vykazují podobnou charakteristiku jako dva takty předchozí. Ačkoli se v případě konsonantického shluku [dv] nejedná o hlásky s formantovou strukturou, jejich plná výslovnost bez známek desonorizace zřejmě také vedla respondenty k vnímání taktu a jeho obrazu jako rytmicky nestejných zvuků. V oscilogramu je i v rámci nízkofrekvenčního obrazu taktu propad mezi intenzitními vrcholy, který tvoří hláska [v], znatelný (obr. 26). Předcházející část taktu však bez segmentálního obsahu mohla percepčně působit spíše jako dlouhá samohláska.



Obr. 26. Oscilogramy taktu [na dva] s jeho nízkofrekvenčním obrazem.

V kapitole 4.4.3.2.1 byly pojednány položky, které byly hodnoceny jako rytmicky téměř či zcela odlišné. Z deseti takových položek jich tři obsahovaly nízkofrekvenční obraz taktu [ɲimɪ]. V rámci položek obsahujících mluvní takt s jeho vlastním obrazem vykazovala také položka N11 s taktem [ɲimɪ] a jeho vlastním nízkofrekvenčním obrazem (obr. 27) medián hodnocení 2, tedy rytmicky podobné, ale ne identické. Důvod byl pravděpodobně týž jako v předchozích případech se stejným taktem: bez segmentálního obsahu je v taktu [ɲimɪ] pouze malý rozdíl v intenzitě sonorních vrcholů.



Obr. 27. Oscilogramy taktu [ɲimɪ] s jeho nízkofrekvenčním obrazem.

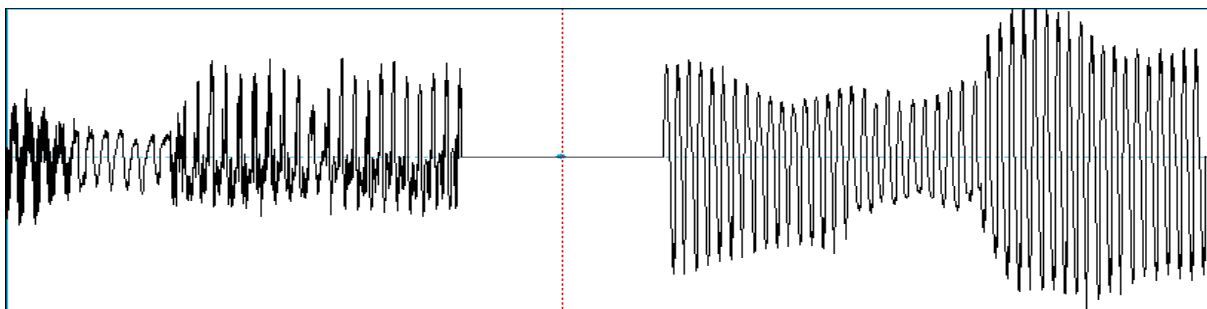
Dalším zajímavým faktem je skutečnost, že hodnocení této položky respondenty mělo dva mody, které spolu nejenže nesousedily, ale dokonce stály na opačných bodech škály – nejčastěji byla tedy tato položka ohodnocena stupni 1 a 5. Respondenti, kteří označili dvojici zvuků v položce za percepčně stejnou, se mohli řídit melodickým průběhem i trváním; nelze vyloučit ani postřehnutí drobného rozdílu v intenzitě. Je možné, že posluchači, kteří označili takty za rytmicky zcela jiné, nízkofrekvenční obraz vnímali jako jakousi nečleněnou pseudoslabiku, čemuž mohlo napomocť i krátké trvání taktu.

#### **4.4.3.3 Položky s hodnocením obsahujícím více modů**

Ačkoli hodnota mediánu hodnocení některých položek žádný extrém nenaznačovala, bližší zkoumání odhalilo, že se jejich hodnocení jednotlivými uživateli zásadně lišila. Kromě výše uvedené položky N11, jejíž hodnocení se vyznačovalo vyšším mediánem i dvěma nesousedícími mody, se v testu vyskytly další tři položky s hodnocením obsahujícím více modů, které spolu nesousedily. Tato charakteristika hodnocení poukazovala na určitou „rytmickou kontroverzi“ – dvě největší skupiny respondentů de facto hodnotily zvuky v položce jako rytmicky podobné a rytmicky odlišné.

Položka P27 s mediánem hodnocení 2 obsahovala dva mody hodnocení: 12 respondentů zvuky hodnotila jako rytmicky téměř identické (hodnocení reprezentováno číslem 1) a jiných 12 posluchačů mluvním taktům v položce přisoudilo hodnocení 4, tedy rytmicky spíše odlišné. Položka obsahovala mluvní takt [zboru] a nízkofrekvenční obrazu taktu [jedna] (obr. 28).



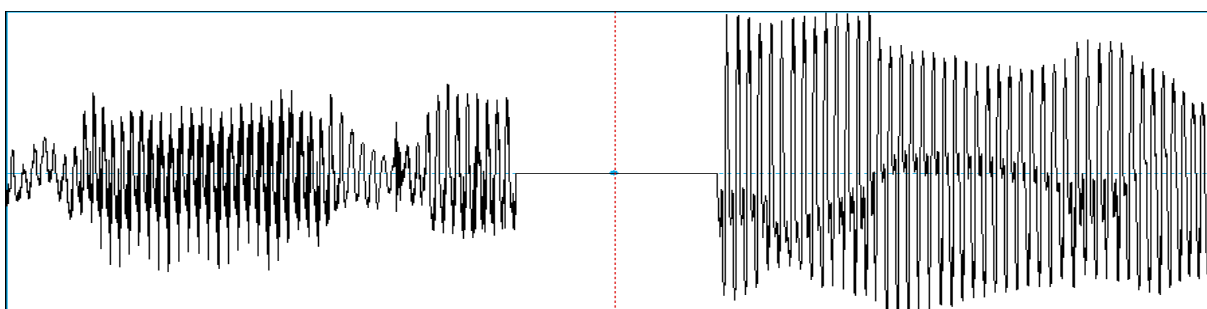


Obr. 28. Oscilogramy taktu [zboru] a nízkofrekvenčního obrazu taktu [jedna].

K hodnocení zvuků jako rytmicky odlišných mohl přispět intonační průběh – zatímco kadence prvního taktu vykazovala mírně stoupavou tendenci (nukleus druhé slabiky byl položen o 1,88 půltónu výš než jádro slabiky první), melodie nízkofrekvenčního obrazu byla monotónní (rozdíl v intonaci jader slabik byl pouze 0,24 půltónu). Také rozdíl v trvání obou zvuků byl poměrně vysoký: dosahoval více než 20 % trvání prvního taktu.

K hodnocení taktů jako rytmicky velmi podobných mohl druhou skupinu posluchačů vést fakt, že oba zvuky obsahovaly percepčně zřetelné dva intenzitní vrcholy. Při bližším pohledu je také možné se domnívat, že svou roli sehrály i trvací proporce zvuků: V taktu [zboru] jsou intenzitní vrcholy, které trvají cca 50 a 60 ms odděleny 40 ms zvuku s nižší intenzitou a v nízkofrekvenčním obrazu taktu [jedna] jsou časové charakteristiky proporčně podobné – trvání intenzitních vrcholů činí cca 100 a 120 ms a pasáž s nižší intenzitou tvoří 80 ms, proporčně tedy jde o dvojnásobky trvání intenzitních vrcholů a předělu v prvním taktu.

Zvuky v položce P33 hodnotilo celkem 14 respondentů jako spíše rytmicky podobné (tedy stupněm 2 na škále subjektivní rytmické podobnosti) a dalších 14 posluchačů jako rytmicky zcela odlišné (stupněm 5). Medián hodnocení položky byl roven hodnotě 2,5. Položka obsahovala plně proslovený mluvní takt [vla:je] a nízkofrekvenční obraz taktu [jme:no] (obr. 29).

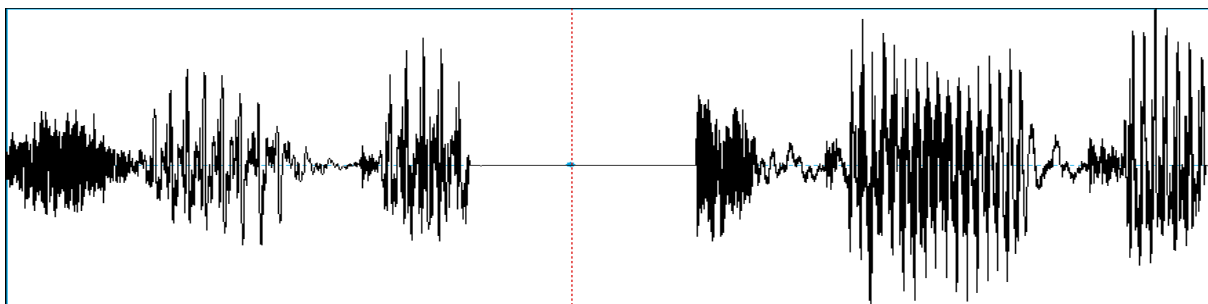


Obr. 29. Oscilogramy taktu [vla:je] a nízkofrekvenčního obrazu taktu [jme:no].

Trvání zvuků v položce se lišilo pouze zanedbatelně (v řádu jednotek milisekund i procent trvání prvního zvuku) a fonotaktická stavba obou taktů byla velice podobná (struktura CCVCV; všechny souhlásky znělé): oba tyto faktory mohly přispět k hodnocení zvuků velkou částí uživatelů jako rytmicky podobných.

Posluchači, kteří hodnotili takty jako odlišné, se pravděpodobně řídili průběhem intonační linie. Kadence prvního taktu byla monotónní (rozdíl v melodii mezi středy nukleů první a druhé slabiky činil pouze 0,19 půltónu), ale na konci nízkofrekvenčního obrazu druhého taktu se objevilo stoupnutí intonace (druhá slabika druhého taktu byla položena o 1,54 půltónu výš než první). Jak již bylo zmíněno v kapitole 2, proměnlivý průběh či vyšší základní frekvence přispívá k vnímání delšího trvání zvuku než u identických úseků s vyrovnanou  $f_0$  (Donovan & Darwin, 1979; Brigner, 1988; Cumming, 2011; Šimko *et al.*, 2015; Dawson *et al.*, 2017). Tuto hypotézu potvrzuje i hodnocení dvojice taktů v položce P173, která obsahuje mluvní takt [jmɛ: no] a nízkofrekvenční obraz taktu [vla: ɣɛ] s manuálně vyhlazeným intonačním průběhem, jako rytmicky velmi podobných. Jak již bylo uvedeno výše, tato položka se vyznačovala mediánem i modem hodnocení 1, zvuky v ní obsažené tedy byly respondenty považovány za rytmicky téměř nebo zcela identické. Položka P168, která měla stejný obsah jako P33 a lišila se od ní jen manuální úpravou melodie, tedy takt [vla: ɣɛ] a nízkofrekvenční obraz taktu [jmɛ: no], při hodnocení na vyhlazení intonační linie mohla paradoxně „doplatit“ – manuálním vyhlazením melodického průběhu zanikl percepční předěl mezi intenzitními vrcholy v nízkofrekvenčním obrazu taktu [jmɛ: no], a pravděpodobně proto měl medián hodnocení posluchačů hodnotu 3 z pětistupňové škály a modus hodnotu 2.

Položka P93 obsahovala dva plně proslovené mluvní takty, a to [sôu to] a [sta: ti] (obr. 30). Hodnota mediánu hodnocení položky posluchači byla vyšší než u předchozích dvou položek, a to 3,5. Dalším rozdílem mezi touto a předchozími dvěma položkami byl fakt, že hodnocení mělo mody dokonce tři: po 10 posluchačích udělilo taktům v položce hodnocení 1, 4 a 5.



Obr. 30. Oscilogramy taktů [sôu to] a [sta:ti].

Podobné složení taktů měla položka P11 s mediánem hodnocení 1,5 a modem 1 – tato tedy byla respondenty hodnocena jako obsahující rytmicky téměř shodné zvuky. Na rozdíl od položky P93 však obsahovala nízkofrekvenční obraz taktu [sta:ti] místo plně prosloveného mluvního taktu. Při zachování segmentálního obsahu pravděpodobně více vynikl rozdíl mezi diftongem a dlouhým vokálem a taktéž byla percepčně vnímatelná odlišnost mezi samostatným konsonantem a souhláskovým shlukem, kterýžto rozdíl v případě porovnání taktu s nízkofrekvenčním obrazem s odstraněnými vysokými frekvencemi šumu zaniká.

## 5 Diskuse

Hlavním výzkumným úkolem předkládané disertační práce bylo vysledovat vztahy mezi fonotaktickou strukturou českého mluvního taktu a řečovým rytmem v češtině. Jelikož je rytmus jevem, který má svůj základ v percepci posluchače (Lehiste, 1977, 1979; Morton *et al.*, 1976; Fletcher, 2010), byl za účelem zjištění tohoto vztahu navržen a uskutečněn percepční experiment. Šturm a Volín (2016) ukázali, že fonotaktická stavba slabiky hraje nezanedbatelnou roli při vnímání rytmu češtiny, a percepční experiment v předkládané práci měl stanovit míru vlivu jednotlivých fonotaktických rysů na vnímání řečového rytmu. Položky v testu sestávaly ze dvou různých dvouslabičných mluvních taktů, či z mluvního taktu a nízkofrekvenčního obrazu jiného či stejného taktu, které se ve většině případů odlišovaly právě jedním hláskovým rysem. Fonotaktické skelety mluvních taktů byly vybrány na základě analýzy četnosti těchto taktů v češtině (Churaňová, 2012). Za činitele, které by mohly ovlivňovat percepci rytmu, byly považovány následující vlastnosti: počet konsonantů v souhláskovém shluku, pozice souhláskové skupiny v mluvním taktu, přítomnost, či absence dlouhé samohlásky ve dvojici taktů a rozdíl v sonoritě či znělosti souhlásek. Posluchači měli za úkol na škále od jedné do pěti určit, nakolik si byly tyto dva zvuky rytmicky podobné.

Z vyhodnocení výsledků experimentu vyplývá, že nejvýznamnějším faktorem pro hodnocení rytmické podobnosti byla **pozice souhláskového shluku v mluvním taktu**. Položky, jejichž takty se odlišovaly právě touto vlastností (např. dvojice [dobjɛ]–[zboru]), byly posluchači hodnoceny jako rytmicky méně podobné než položky, v nichž se tento rozdíl neobjevil. Tato tendence se projevila v analýze celého souboru položek ( $t(221) = -5,96; p < 0,001$ ), i při zahrnutí pouze těch položek, které neobsahovaly mluvní takt s jeho vlastním nízkofrekvenčním stínem:  $t(182) = -4,54; p < 0,001$ . Pro významnost tohoto rysu pro vnímání rytmické podobnosti svědčí i položková analýza, v jejímž rámci byly zkoumány takové položky, které vykazovaly extrémní mediány hodnocení. Jako rytmicky téměř či zcela nepodobné bylo posluchači hodnoceno celkem deset položek, z nichž čtyři obsahovaly rozdíl v pozici souhláskového shluku v rámci mluvních taktů; rozdíly ve všech ostatních rysech byly méně četné. V případě opačném, tedy v analýze dvojic hodnocených jako rytmicky téměř identických, rozdíl v pozici souhláskového shluku nefiguroval vůbec, což podporuje závěry vzniklé ze souhrnných analýz. Významnost tohoto rysu pro vnímání rytmické podobnosti byla pravděpodobně

způsobena značnými rozdíly v trvání úseků mezi intenzitními vrcholy první a druhé slabiky u obou taktů: jen u zmíněných čtyř extrémně hodnocených položek činily tyto rozdíly 50 a více milisekund.

Dalším relativně významným rysem pro vnímání rytmické podobnosti se ukázal **počet souhlásek v konsonantickém shluku** (např. dvojice [sôu to]–[sta:ti]). Rozdíl v hodnocení byl významný jen tehdy, pokud analýza zahrnovala všechny položky, tedy i ty, které obsahovaly takt s jeho vlastním nízkofrekvenčním obrazem:  $t(221) = -2,56$ ;  $p < 0,05$ ; v případě vynechání těchto položek se ukázal nevýznamný trend podobného směru:  $t(182) = -0,5$ ;  $p = 0,617$ . Zatímco rys pozice konsonantického shluku v taktu zůstal silným i v porovnání s pouze těmi položkami, které vždy obsahovaly nějaký další rozdíl ve fonotaktické struktuře taktů, rys počtu souhlásek v konsonantickém shluku v rámci mluvního taktu po odstranění části vzorku s takty bez jakýchkoliv rozdílů ztratil na významnosti. Dvě položky, které obsahovaly rozdíl v tomto rysu a byly hodnoceny jako rytmicky téměř či zcela odlišné, zároveň vykazovaly i související odlišnosti, jako je vložení epentetického schwa mezi souhlásky ve skupině či velký rozdíl v trvání celého taktu.

Lze konstatovat, že oba výše uvedené faktory, tedy pozice konsonantů ve vyšší prozodické jednotce a počet souhlásek ve shluku hrají nikoli bezvýznamnou roli při vnímání rytmu. Konsonantické shluky v souvislosti s rytmem řeči zmínila např. Dauerová (1983), která přítomnost souhláskových skupin spojovala spíše s taktově izochronními jazyky. Jak už bylo popsáno v úvodu, čeština není jednoznačně zařaditelná do skupiny taktově, či slabičně izochronních jazyků (Duběda, 2004; Dankovičová & Dellwo, 2007); jisté však je, že více či méně složité souhláskové skupiny, které se v češtině vyskytují, přispívají k percepci jejího rytmu.

Důležitou roli ve vnímání rytmické podobnosti hraje také přítomnost **fonologicky dlouhé samohlásky** v mluvním taktu (např. dvojice [vi:ʃtɛ]–[mi:sti]). Podle výsledků byly mluvní takty, které obsahovaly výhradně krátké samohlásky, hodnoceny jako méně si rytmicky podobné než takty, v nichž dlouhý vokál figuroval, což se potvrdilo v analýze všech položek ( $t(221) = -2,55$ ;  $p < 0,05$ ) i ve skupině položek bez zahrnutí dvojic s identickými takty:  $t(182) = -2,72$ ;  $p < 0,01$ . Také v rámci položkové analýzy se ukázala přítomnost dlouhého vokálu ve dvojici taktů jako důležitá: z 11 položek hodnocených jako rytmicky téměř stejných jich dlouhou samohlásku obsahovalo šest, a to i přes

skutečnost, že položky s fonologicky dlouhými samohláskami tvořily pouze necelých 31 % ze všech položek testu; zbylých 69 % připadlo na položky s výhradně krátkými vokály. Že byly položky s dlouhými vokály hodnoceny jako podobnější, podporuje i fakt, že mezi 10 dvojicemi taktů, které posluchači označili jako rytmicky velmi odlišné, byla pouze jediná dvojice obsahující dlouhou samohlásku. Všechny tyto závěry shodně naznačují, že přítomnost dlouhého vokálu je pro hodnocení rytmu mluvního taktu posluchačem velmi výrazným vodítkem.

Jako méně podstatný rys při posuzování rytmické podobnosti se ukázala přítomnost **rozdílu v sonoritě znělých hlásek** ve dvojici taktů. Podle souhrnné analýzy byly položky obsahující znělé souhlásky odlišné sonority (např. [vi:ʃtɕe] a [mi:stɪ]) hodnoceny jako méně rytmicky si podobné než ostatní položky:  $t(173) = -1,17$ ;  $p \doteq 0,244$ . Po odstranění položek s identickými takty se objevila statisticky významná opačná tendence – položky lišící se v sonoritě znělých konsonantů byly tedy hodnoceny jako rytmicky méně rozdílné než jiné:  $t(134) = 2,64$ ;  $p < 0,01$ . Tento výsledek je však pravděpodobně zapříčiněn návrhem experimentu: po odstranění položek se stejnými mluvními takty zůstaly k porovnání proti vlivu sonority pouze položky s rozdíly v jiných vlastnostech, často silnějších než rozdíl v sonoritě souhlásek (např. počet konsonantů ve skupině, pozice souhláskového shluku). Analýza položek, které byly hodnoceny jako rytmicky podobné, souhlasí s obecnými trendy: osm z celkových 11 položek se vyznačovalo rozdílem právě v tomto rysu. Sonorita se ovšem objevila i ve třech případech dvojic taktů, které byly posluchači označeny za rytmicky zcela odlišné. U těchto položek se ale projevy i další odlišnosti, které pravděpodobně extrémní hodnocení zapříčinilo (např. vložený vokalický prvek, nejasný počet slabik kvůli sonornímu průběhu celého slova).

V provedeném experimentu nebyl zaznamenán žádný významný vliv **rozdílu ve znělosti hlásek** na vnímání rytmické podobnosti (např. dvojice [mezi]–[jako]). Trendy byly neprůkazné u analýzy všech položek ( $t(221) = 1,31$ ;  $p \doteq 0,192$ ) i s odstraněním položek s identickými takty ( $t(182) = -0,88$ ;  $p \doteq 0,381$ ). V rámci analýzy dvojic taktů hodnocených jako rytmicky ekvivalentní se znělost objevila dvakrát z celkových 11 případů. I zde je namísto předpokládat, že rys znělosti byl oproti ostatním vlastnostem příliš subtilní na to, aby se projevil. Jak už také bylo poznamenáno v kapitole 4.2.5, fonologická znělost nemusí být signalizována vždy jen přítomností

základní frekvence; hláska může být vnímána jako znělá např. díky temporálním vlastnostem (Kent & Read, 2002), nebo i frekvenci základního tónu (Hogan & Rozsypal, 1980). Je tedy možné se domnívat, že to, jestli posluchač vnímá souhlásku jako znělou, určuje spíše než přítomnost základního tónu znalost daného jazyka. Ačkoli čeština bývá považována za jazyk, pro nějž je kontrast znělosti důležitý, z těchto výsledků je patrné, že znělost nelze prohlásit za robustní parametr pro určení rytmu řeči.

Kromě výše uvedených fonotaktických či fonologických vlivů na vnímání rytmu byly zkoumány i dva faktory, které rovněž mohly přispět k percepci rytmické podobnosti a které se nacházely na vyšších rovinách než segmentální. První takovou vlastností byla **přítomnost manuálně vyhlazené intonační linie**. T-testy neodhalily žádný průkazný vliv tohoto faktoru na percepci taktů jako rytmicky si podobných, nebo odlišných. Tento výsledek mohl mít příčinu v návrhu experimentu, který si kladl za cíl přijímat do výběru položek pouze mluvní takty, na nichž se nerealizovala melodická kadence. Zpravidla tedy měly takty neutrální intonaci již přirozeně. V konkrétních případech, kde tomu tak nebylo, se ale vliv vyhlazení intonační linie projevit mohl, jak ukazovalo například rozdílné hodnocení položek P33 a P173, které obsahovaly stejné takty, ale jen jedna z položek měla manuálně vyhlazenou intonační kadenci (viz kapitola 4.4.3.3).

Dalším zkoumaným faktorem, tentokrát biologickým, byl **rozdíl v pohlaví mluvčích** jednotlivých taktů v položce. Statisticky významný rozdíl se objevil pouze v souhrnné analýze všech položek:  $t(221) = 3,14$ ;  $p < 0,01$ , přičemž jako rytmicky podobnější byly hodnoceny položky, u nichž se pohlaví mluvčích shodovalo. Na souboru položek bez identických taktů vyšel trend stejného směru, ale zcela nevýznamný:  $t(182) = 0,16$ ;  $p \doteq 0,876$ . Je pravděpodobné, že v tomto případě výsledky spíše ukazují na důležitost individuality mluvčího než pohlaví jakožto podstatného faktoru v percepci rytmické podobnosti.

Ačkoli se některé faktory, které by mohly ovlivňovat percepci rytmu řeči, při vyhodnocení výsledků experimentů ukázaly jako téměř či zcela nevýznamné, nelze o nich jednoznačně prohlásit, že se na vnímání rytmické podobnosti vůbec nepodílejí. Jak již bylo zmíněno, kontrast sonority a kontrast znělosti lze považovat za subtilnější rozdíly ve fonotaktické struktuře mluvních taktů, a pokud významnost těchto kontrastů testujeme proti položkám, které obsahují rozdíly v robustnějších faktorech (jako je například pozice konsonantického shluku v rámci taktu), je namístě očekávat, že se

rozdíl ve vnímání taktů lišících se pouze znělostí či sonoritou konsonantu neukáže jako podstatný. Pro vynesení jednoznačného závěru by bylo nutné rozšířit soubor položek tak, aby obsahoval dvojice mluvních taktů s různým segmentálním obsahem, avšak stejných fonotaktických vzorců – porovnávaly by se tedy mluvní takty, které by obsahovaly rozdílná slova, ale jejich konsonanticko-vokalické složení by bylo totožné (např. takty obsahující slova *zlato* a *vrata* oba respektují vzorec ZSVNV). Přidáním těchto položek by ale výrazně vzrostl časový rozsah percepčního testu, který byl i v realizované podobě s 25 minutami trvání a 210 položkami pro posluchače téměř hraniční zátěží.

Zda se faktory, jako je kontrast sonority a kontrast znělosti, skutečně na vnímání řečového rytmu nepodílejí, je možné ověřit v budoucím výzkumu, který by mohl být zaměřen právě na tyto dva rysy. Takový experiment by obsahoval dvojice taktů odpovídající nejčastějším konsonanticko-vokalickým vzorcům, které by se lišily buď pouze v rysu sonority jednoho konsonantu či konsonantické skupiny, nebo výhradně v rysu znělosti jednoho obstruentu či skupiny obstruentů. Experiment by také zahrnoval dvojice mluvních taktů s různým segmentálním obsahem a identickými fonotaktickými vzorci. Pro vyhodnocení by bylo možné porovnávat hodnocení skupiny obsahující specifický rozdíl s hodnocením skupiny bez fonotaktického rozdílu a také by bylo možné určit, který z faktorů (sonorita, nebo znělost), je pro vnímání rytmické podobnosti významnější.

Dalším směrem, kterým je možné se v navazujícím výzkumu vydat, je vnímání rytmické podobnosti jednotlivými posluchači. V rámci položkové analýzy vyšla najevo bimodální hodnocení položek, při jejichž posuzování se zhruba rovnoměrné počty respondentů řídily rozdílnými zvukovými vlastnostmi. Jedna z linií případného návazného výzkumu by tedy mohla zjednodušit skladbu položek, ale věnovat se většímu vzorku populace, který by tento jev pomohl objasnit.



## 6 Závěr

Tato disertační práce měla prozkoumat vztahy mezi fonotaktickou stavbou českého mluvního taktu a řečovým rytmem. Mluvní takt byl pro analýzu zvolen proto, že právě on je základní jednotkou, na které se realizuje rytmus řeči; pro účely této práce byly vybrány takty, jež odpovídaly třem nejčastějším dvouslabičným konsonanticko-vokalickým vzorcům v češtině: CVCV, CCVCV a CVCCV. V percepčním experimentu, který zahrnoval srovnání plně proslovených taktů a jejich nízkofrekvenčních obrazů, posluchači určovali, nakolik jsou si takty různých i stejných vzorců rytmicky podobné.

Výsledky ukázaly, že nejsilnější vliv na vnímání rytmické podobnosti má z fonotaktických faktorů pozice souhláskového shluku v mluvním taktu. Pokud se porovnávané jednotky v rámci položky odlišovaly právě touto vlastností (tzn. šlo o vzájemné srovnání taktů s konsonanticko-vokalickými vzorci CCVCV a CVCCV), byly tyto takty posluchači konzistentně hodnoceny jako rytmicky spíše nepodobné. O něco méně silným, ale stále relativně robustním faktorem pro percepci řečového rytmu byl také počet souhlásek v konsonantickém shluku – takt odpovídající vzorci CVCV byl posluchači vnímán jako rytmicky odlišný od jednotek respektujících vzorec CCVCV nebo CVCCV. Oproti tomu za rytmicky si navzájem podobnější mluvní takty označovali respondenti takové položky, v nichž se v obou taktech na stejném místě vyskytovala dlouhá samohláska, na rozdíl od taktů pouze s krátkými vokály; i tento trend dosáhl ve vyhodnocení experimentu statistické významnosti.

Ačkoli faktory sonority a znělosti souhlásek (tzn. porovnávané takty se lišily pouze rysem sonority u znělých souhlásek, nebo výhradně rysem znělosti u obstruentů) měly jen minimální nebo neprůkazný vliv na percepci rytmické podobnosti mluvních taktů, nelze o nich bez dalšího přezkoumání prohlásit, že k vnímání rytmu nepřispívají. Způsob, jakým byl experiment navržen, zachycoval jen nejsilnější trendy ve vnímání rytmu řeči ve vztahu k fonotaktické struktuře mluvních taktů, protože se zpravidla testovaly průměry hodnocení položek lišících se jedním rysem proti průměrům hodnocení položek lišících se rysem jiným, často silnějším (např. pozice souhláskové skupiny v rámci mluvního taktu). Pro zachycení případného vlivu těchto subtilnějších faktorů by bylo nutné provést podobný experiment, tentokrát ale zaměřený pouze na uvedené rysy.

V rámci položkové analýzy, která následovala po souhrnných analýzách výsledků percepčního experimentu, byla věnována pozornost těm taktům, které byly respondenty konzistentně hodnoceny jako rytmicky téměř či zcela stejné (11 položek), a také dvojicím, které naopak posluchači považovali za rytmicky výrazně či úplně odlišné (10 položek). Také byly prozkoumány položky, které obsahovaly plně proslovený takt s jeho vlastním nízkofrekvenčním stínem, ale zároveň dvojice taktů nebyly hodnoceny jako zcela či téměř stejné (6 položek). Nakonec byly prozkoumány položky, jejichž hodnocení mělo více nesousedících modů, z nichž alespoň jeden byl krajním stupněm hodnotící škály (4 položky). Položková analýza byla vesměs v souladu s trendy určenými souhrnnými výsledky. V případech, kdy tomu tak nebylo, byl obvykle zřetelně identifikován prvek, který extrémní hodnocení zapříčinil (např. vložený slabičný prvek, rozdíl v intonační kadenci taktů apod.).

Předkládaná disertační práce si nekladla za cíl podat vyčerpávající výčet všech faktorů, které mohou ovlivňovat vnímání rytmu řeči v češtině, nýbrž podchytit nejsilnější vlivy, které se v tomto vztahu vyskytují. Vyzkoumané závěry tedy umožňují základní vhled do tendencí, které jsou spojeny s vnímáním rytmické podobnosti ve vztahu k fonotaktické struktuře základních prozodických jednotek v češtině, a poskytují vodítka pro navazující studie. Budoucí výzkum by mohl zahrnovat srovnatelné experimenty, které by rozšířily testovaný materiál, věnovaly větší prostor subtilnějším rysům, jako je znělost a sonorita, zohledňovaly takty delší než dvouslabičné, případně i celé promluvové úseky, a byly by otestovány na větším vzorku posluchačů, což rozsah tohoto výzkumu plně neumožňoval.

## Literatura

- Abercrombie, D. (1964). Syllable quantity and enclitics in English. In: Abercrombie, D., Fry, D. B., MacCarthy, P. A. D., Scott, N. C. & Trim, J. L. M. (Eds.), *In Honour of Daniel Jones*, 216–222. London: Longman.
- Abercrombie, D. (1967). *Elements of general phonetics*. Edinburgh: Edinburgh University Press.
- Balkó, I. (1999). K fonetickému výzkumu tempa řeči a tempa artikulace v čteném textu a spontánním projevu. In: Čechová, M. & Moldanová, D. (Eds.), *Jinakost, cizost v jazyce a literatuře. Sborník z mezinárodní konference*, 38–44. Ústí nad Labem: Univerzita Jana Evangelisty Purkyně.
- Balkó, I. (2001). K některým příčinám variability tempa řeči a tempa artikulace televizních moderátorů v začáteční, střední a závěrečné fázi relace o počasí. In: Moldanová, D. & Čechová, M. (Eds.), *Konec a začátek v jazyce a literatuře. Sborník z mezinárodní konference*, 214–218. Ústí nad Labem: Univerzita Jana Evangelisty Purkyně.
- Balkó, I. (2003). *Tempo artikulace a tempo řeči v různých řečových úlohách*. Disertační práce. Praha: Univerzita Karlova.
- Balkó, I. (2005). K výzkumu tempa řeči a tempa artikulace v různých řečových úlohách. „*Bohemistika*“ 2005, 3, 185–198.
- Barbosa, P. A., Arantes, P., Meireles, A. R. & Vieira, J. M. (2005). Abstractness in speechmetronome synchronization: p-centres as cyclic attractors. In: *Proceedings of Interspeech 2005*, 1441–1444. Lisbon: ISCA.
- Bartoň, T., Cvrček, V., Čermák, F., Jelínek, T. & Petkevič, V. (2009). *Statistiky češtiny*. Praha: Nakladatelství Lidové noviny.
- Bartošek, J. (2000). Mluvní tempo v rozhlasu a v televizi. In: *Psychotrofon*, 1, s. 78–82. Olomouc: Univerzita Palackého.
- Bartošová, P. (2016). *Tempo řeči v jevištní češtině ve dvou obdobích s větším časovým odstupem*. Diplomová práce. Praha: Univerzita Karlova.
- Beckman, M. E. (1986). *Stress and Non-Stress Accent*. Dordrecht: Foris.

- Bičan, A. (2011). *Phonotactics of Czech*. Disertační práce. Brno: Masarykova univerzita.
- Bičan, A. (2015). Kvantitativní analýza slabiky v českém lexikonu. *Linguistica Brunensia* 63(2), 87–107. Brno: Masarykova univerzita.
- Blevins, J. (1995). The Syllable in Phonological Theory. In: Goldsmith, J. (Ed.), *Handbook of phonological theory*, 206–244. London: Basil Blackwell.
- Boersma, P. & Weenink, D. (2018). *Praat: doing phonetics by computer* [počítačový program], verze 6.0.30. Získáno z <http://www.praat.org/>.
- Borovičková, B. & Maláč, M. (1967). *The Spectral Analysis of Czech Sound Combinations*. Praha: Academia.
- Brigner, W. L. (1988). Perceived duration as a function of pitch. *Perceptual and Motor Skills*, 1988, 67, 301–302.
- Buxton, H. (1983). Temporal predictability in the perception of English speech. In: Cutler, A. & Ladd, D. R. (Eds.), *Prosody: Models and Measurements*, 111–121. Berlin: Springer-Verlag.
- Castleman, W. A. & Diehl, R. L. (1996). Effects of fundamental frequency on medial and final [voice] judgments. *Journal of Phonetics*, 24(4), 383–398.
- Clements, G. N. (1990). The role of the sonority cycle in core syllabification. In: Kingston, J. & Beckman, M. E. (Eds.), *Papers in laboratory phonology I: Between the grammar and the physics of speech*, 283–333. Cambridge: Cambridge University Press.
- Cooper, A. M., Whalen, D. H. & Fowler, C. A. (1986). P-centers are unaffected by phonetic categorization. *Perception & Psychophysics*, 39, 187–196.
- Crystal, T. H. & House, A. (1988). Segmental durations in connected speech signals: Syllabic stress. *Journal of the Acoustical Society of America*, 83, 1574–1585.
- Cumming, R. (2011). The effect of dynamic fundamental frequency on the perception of duration. *Journal of Phonetics*, 39, 375–387.
- Cummins, F. (2009). Rhythm as an affordance for the entrainment of movement. *Phonetica*, 66, 15–28.
- Cutler, A. (1981). The reliability of speech error data. *Linguistic*, 19, 561–582.
- Daneš, F. (1957). *Intonace a věta ve spisovné češtině*. Praha: Academia.

- Daneš, F. Intonace v textu (promluvě) (1985). *Slovo a slovesnost*, 46, 83–100.
- Dankovičová, J. (1997). The domain of articulation rate in Czech. *Journal of Phonetics*, 25, 287–312.
- Dankovičová, J. (1998). *The linguistic basis of articulation rate variation in Czech*. Disertační práce. Oxford: Oxford University.
- Dankovičová, J. & Dellwo, V. (2007). Czech speech rhythm and the rhythm class hypothesis. In: *Proceedings of the XVIth ICPHS*, 1241–1244. Saarbrücken: Organizing Committee.
- Dauer, R. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of the International Phonetic Association*, 11, 51–62.
- Dawson, C., Aalto, D., Šimko, J. & Vainio, M. (2017). The influence of fundamental frequency on perceived duration in spectrally comparable sounds. *PeerJ*, 5, e3734.
- Dellwo, V., Fourcin, A. & Abberton, E. (2007). Rhythmical classification of languages based on voice parameters. In: *Proceedings of the XVIth ICPHS*, 1129–1132. Saarbrücken: Organizing Committee.
- Donovan, A. & Darwin, C. J. (1979). The perceived rhythm of speech. In: Fischer-Jørgensen, E. *et al.* (Eds.), *Proceedings of IXth International Congress of Phonetic Sciences*, 268–274. Copenhagen: University of Copenhagen.
- Duběda, T. (2004). K izosylabičnosti a izochronnosti v češtině. In: *Sborník z Konference česko-slovenské pobočky ISPhS 2004*, 19–28. Praha: Univerzita Karlova.
- Duběda, T. (2005). *Jazyky a jejich zvuky. Univerzálie a typologie ve fonetice a fonologii*. Praha: Karolinum.
- Fletcher, J. (2010). The Prosody of Speech: Timing and Rhythm. In: Hardcastle, W., Laver, J. & Gibbon, F. (Eds.), *The Handbook of Phonetic Sciences*, 523–602. Spojené království: Wiley-Blackwell Publishing.
- Fry, D. B. (1955). Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America*, 27, 765–768.
- Fry, D. B. (1958). Experiments in the perception of stress. *Language and Speech*, 1, 126–152.

- Ghitza, O. & Greenberg, S. (2009). On the possible role of brain rhythms in speech perception: Intelligibility of time compressed speech with periodic and aperiodic insertions of silence. *Phonetica*, 66, 113–126.
- Goldsmith, J. A. (2011). Syllables. In: Goldsmith, J. A., Riggle, J. & Yu, A. C. L. (Eds.), *The handbook of phonological theory*, 164–196. Oxford: Wiley-Blackwell.
- Grabe, E. & Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. In: Warner, N. & Gussenhoven, C. (Eds.), *Papers in laboratory phonology* 7, 515–546. Berlin: Mouton de Gruyter.
- Grossberg, S. (2003). Resonant neural dynamics of speech perception. *Journal of Phonetics*, 31(3–4), 423–445.
- Hála, B. (1948). *Úvod do fonetiky*. Praha: Melantrich.
- Hála, B. (1956). *Slabika, její podstata a vývoj*. Praha: Československá Akademie věd.
- Hála, B. (1962). *Uvedení do fonetiky češtiny na obecně fonetickém základě*. Praha: Nakladatelství Československé akademie věd.
- Hála, B. (1975). *Fonetika v teorii a praxi*. Praha: Státní pedagogické nakladatelství.
- Hála, B. & Sovák, M. (1947). *Hlas, řeč, sluch*. Praha: Česká grafická unie.
- Handbook of the International Phonetic Association: a guide to the use of the International Phonetic Alphabet* (1999). Cambridge: Cambridge University Press.
- Havránek, B. & Jedlička, A. (1988). *Česká mluvnice*. Praha: Státní pedagogické nakladatelství.
- Heldner, M. & Strangert, E. (2001). Temporal effects of focus in Swedish. *Journal of Phonetics*, 29, 329–361.
- Hermes, D. J. (2006). Stylization of pitch contours. In: Sudhoff *et al.* (Eds.), *Methods in empirical prosody research*, 29–61. Berlin: Walter de Gruyter.
- Hirst, D. & Di Cristo, A. (1998). A survey of intonation systems. In: Hirst, D. & di Cristo, A. (Eds.), *Intonation systems: a survey of twenty languages*, 1–43. Cambridge: Cambridge University Press.

- Hogan, J. T. & Rozsypal, A. J. (1980). Evaluation of vowel duration as a cue for the voicing distinction in the following word-final consonant. *Journal of the Acoustical Society of America*, 67, 1764–1771.
- Huggins, A. W. F. (1979). Some effects on intelligibility of inappropriate temporal relations within speech units. In: Fischer-Jørgensen, E. *et al.* (Eds.), *Proceedings of IXth International Congress of Phonetic Sciences*, 283–289. Copenhagen: University of Copenhagen.
- Chen, F. R. (1980). *Acoustic characteristics and intelligibility of clear and conversational speech at the segmental level*. Diplomová práce. Cambridge, MA: MIT.
- Chlumský, J. (1911). *Pokus o měření českých zvuků a slabik v řeči souvislé*. Praha: Česká akademie.
- Chlumský, J. (1928). *Česká kvantita, melodie a přízvuk*. Praha: Česká akademie věd a umění.
- Churaňová, E. (2012). *Fonotaktická osnova českého slova a mluvního taktu*. Diplomová práce. Praha: Univerzita Karlova.
- Churaňová, E. (2013). The consonantal-vocalic structure of the Czech word and stress group. *AUC Philologica 1/2014, Phonetica Pragensia XIII*, 79–90.
- Churaňová, E. (2015). Speaker, listener and speech-metronome synchronization. In: Niebuhr, O. & Skarnitzl, R. (Eds.), *Tackling the Complexity in Speech*, 27–37. Praha: FF UK.
- Churaňová, E., Šturm, P. & Weingartová, L. (2015). Changes in segmental timing in slow and fast metronome-synchronized speech. In: *Proceedings of the 18<sup>th</sup> ICPHS*. Glasgow: University of Glasgow.
- Janota, P. & Palková, Z. (1974). Auditory evaluation of stress under the influence of context. *AUC Philologica 2/1974, Phonetica Pragensia*, 4, 29–59.
- Jantzen, K. J., Steinberg, F. L. & Kelso, J. A. (2002). Practice-dependent modulation of neural activity during human sensorimotor coordination: A functional magnetic resonance imaging study. *Neuro-science Letters*, 332, 205–209.
- Kaiser, L. (1964). Phonetic similarity apart from linguistic affinity. *Zeitschrift für Phonetik*, 17, 243–249.

- Kelso, J. A. (1995). *Dynamic Patterns: The Self-Organization of Brain and Behavior*. Cambridge: MIT Press.
- Kent, R. D. & Read, C. (2002). *The acoustic analysis of speech*. Australia: Singular/Thomson Learning.
- Klatt, D. H. (1973). Durational characteristics of prestressed word-initial consonant clusters in English. *Research Laboratory of Electronics: MIT Quarterly progress report*, 108, 253–260.
- Klatt, D. H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, 59, 1208–1221.
- Kohler, K. J. (1986). Invariance and variability in speech timing: From utterance to segment in German. In: Perkell, J. & Klatt, D. H. (Eds.), *Invariance and Variability in Speech Processes*, 268–89. Hillsdale, NJ: Lawrence Erlbaum.
- Kohler, K. J. (2009). Rhythm in speech and language. A new research paradigm. *Phonetica*, 66, 29–45.
- Koreman, J. (2006). Perceived speech rate: The effects of articulation rate and speaking style in spontaneous speech. *Journal of the Acoustical Society of America*, 119, 582–596.
- Krčmová, M. (2005). Výslovnost ve veřejném projevu. In: Jelínek, M., Švandová, B. (Eds.), *Argumentace a umění komunikovat*, 297–312. Brno: Masarykova univerzita.
- Kučera, H. (1961). *The phonology of Czech*. The Hague: Mouton.
- Kučera, H. (1963). Mechanical phonemic transcription and phoneme count of Czech. *International Journal of Slavic Linguistics and Poetics*, 6, 36–50.
- Kučera, H. & Monroe, G. K. (1968). *A comparative quantitative phonology of Russian, Czech and German*. New York: American Elsevier Publishing Company.
- Ladefoged, P. (1967). *Three Areas of Experimental Phonetics*. Oxford: Oxford University Press.
- Lehiste, I. (1977). Isochrony reconsidered. *Journal of Phonetics*, 5, 253–263.



- Lehiste, I. (1979). Temporal relations within speech units. In: Fischer-Jørgensen, E. *et al.* (Eds.). *Proceedings of IXth International Congress of Phonetic Sciences*, 241–244. Copenhagen: University of Copenhagen.
- Lisker, L. (1974). On “explaining” vowel duration variation. *Glossa* 8(2), 233–246.
- Low, E. L., Grabe, E. & Nolan, F. (2001). Quantitative characterisations of speech rhythm: Syllable-timing in Singapore English. *Language and Speech*, 43(4), 377–401.
- Ludvíková, M. (1987). Čísla o hláskách. In: Těšitelová, M. *et al.* (Eds.), *O češtině v číslech*, 91–108. Praha: Academia.
- Ludvíková, M. & Kraus, J. (1966). Kvantitativní vlastnosti soustavy českých fonémů. *Slovo a slovesnost*, 27, 334–344.
- Machač, P. & Skarnitzl, R. (2009). *Fonetická segmentace hlásek*. Praha: Nakladatelství Epocha.
- Machač, P. (2006). *Temporální a spektrální struktura českých explozív*. Disertační práce. Praha: Univerzita Karlova.
- Machač, P. (2008). Desonorizace českých intervokalických frikativ. In: Volín, J. & Janoušková, J. (Eds.), *AUC Philologica 2/2007, Phonetica Pragensia XI*, 105–116. Praha: Karolinum.
- MacNeilage, P. F. (1998). The frame/content theory of evolution of speech production. *Behavioral and brain sciences*, 21, 499–546.
- Marcus, S. M. (1981). Acoustic determinants of perceptual center (P-center) location. *Perception & Psychophysics*, 30, 247–256.
- Mazlová, V. (1946). Jak se projevuje zvuková stránka češtiny v hláskových statistikách. *Naše řeč*, 30(6, 7, 8), 101–111, 146–151.
- Morton, J., Marcus, M. & Frankish, C. (1976). Perceptual centres (P-centres). *Psychological Review*, 83, 405–408.
- Müllerová, O. & Nekvapil, J. (1986). Pauzy v mluveném textu. *Slovo a slovesnost*, 47(2), 105–113.
- Nespor, M. (1990). On the rhythm parameter in phonology. In: Roca, I. M. (Ed.), *Logical Issues in Language Acquisition*, 157–175. Dordrecht: Foris.

- Ohala, J. J. (1975). The temporal regulation of speech. In: Fant, G. & Tatham, M. (Eds.), *Auditory Analysis and Perception of the Speech*, 431–454. London: Academic Press.
- Ondráčková, J. (1954). O mluvním rytmu v češtině. *Slovo a slovesnost*, 15, 24–29, 145–157.
- Ondráčková, J. (1962). K analýze přízvuchnosti, zvláště v češtině. *AUC Philologica – Slavica Pragensia IV*, 81–88.
- Ondrušková, L. (2011). *Zvukové vlastnosti jednoslabičných slov v semispontánním dialogu a hlasitém čtení*. Diplomová práce. Praha: Univerzita Karlova.
- Palková, Z. (1994). *Fonetika a fonologie češtiny*. Praha: Karolinum.
- Palková, Z. (2004). Přízvukový takt ve struktuře češtiny. In: Hladká, Z. & Karlík, P. (Eds.), *Čeština – Univezália a specifika 5*, 399–408. Praha: Nakladatelství Lidové noviny.
- Palková, Z. (2006). *Textové dispozice pro členění na intonační fráze v češtině*. In: Janoušková, J. & Palková, Z. (Eds.), *Kapitoly z fonetiky a fonologie slovanských jazyků*, 227–239. Praha: FF UK.
- Palková, Z., Veroňková, J., Volín, J. & Skarnitzl, R. (2004). Stabilizace některých termínů pro fonetický popis češtiny v závislosti na nových výsledcích výzkumu. In: Duběda, T. (Ed.), *Sborník z konference česko-slovenské pobočky ISPhS 2004*, 65–74. Praha: Karolinum.
- Picheny, M. A., Durlach, N. I. & Braidá, L. D. (1986). Speaking clearly for the hard of hearing, II: Acoustic characteristics of clear and conversational speech. *Journal of Speech and Hearing Research*, 29, 434–446.
- Pike, K. L. (1945). *The intonation of American English*. University of Michigan Press: Ann Arbor.
- Pollák, P., Volín, J. & Skarnitzl, R. (2007). HMM-Based Phonetic Segmentation in Praat Environment. In: *Proceedings of the XIIth International Conference "Speech and computer – SPECOM 2007"*, 537–541, Moskva.
- Port, R. F. (2003). Meter and speech. *Journal of Phonetics*, 31, 599–611.
- Port, R. F., Al-Ani, S., Maeda, S. (1980). Temporal compensation and universal phonetics. *Phonetica*, 37, 235–252.

- Psutka, J., Müller, L., Matoušek, J. & Radová, V. (2006). *Mluvíme s počítačem česky*. Praha: Academia.
- Ramus, F., Nespore, M. & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, 73(3), 265–292.
- Rao, M., Harrington, D. L., Haaland, K. Y., Bobholz, J. A., Cox, R. W. & Binder, J. R. (1997). Distributed neural systems underlying the timing of movements. *Journal of Neuroscience*, 17, 5526–5535.
- Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin & Review*, 12(6), 969–992.
- Repp, B. H. & Keller, P. E. (2004). Adaptation to tempo changes in sensorimotor synchronization: Effects of intention, attention, and awareness. *Quarterly Journal of Experimental Psychology*, 57 A, 499–521.
- Rietveld, T., Kerkhoff, J. & Gussenhoven, G. (2004). Word prosodic structure and vowel duration in Dutch. *Journal of Phonetics*, 32, 349–371.
- Rochet-Capellan, A. & Schwartz, J. (2005). The labial-coronal effect and CVCV stability during reiterant speech production: an articulatory analysis. In: *Proceedings of the 9th International Conference on Speech Communication and Technology (InterSpeech)*, 1013–1016. Lisbon: ISCA.
- Romportl, M. (1951). *K tónovému průběhu v mluvené češtině*. Praha: KČSN.
- Romportl, M. (1958). *Zvuková stránka souvislé řeči v nářečích na Těšínsku: Fonetická studie*. Ostrava: Krajské nakladatelství.
- Romportl, M. (1985). *Základy fonetiky*. Praha: Karolinum.
- Rousset, E. (2004). *Structures syllabiques et lexicales dans les langues du monde*. Disertační práce. Grenoble: ICP.
- Rubovičová, C. (2014). *Tempo řeči a realizace pauz při konsektivním tlumočení do češtiny ve srovnání s původními českými projevy*. Diplomová práce. Praha: Univerzita Karlova.
- Sedláková, J. (1989). *Proměnlivost individuálního tempa řeči v mluvených projevech monologického charakteru*. Diplomová práce. Praha: Univerzita Karlova.

- Skarnitzl, R. (2010). Prague Phonetic Corpus: status report. *AUC Philologica 1/2009, Phonetica Pragensia XII*, 65–67.
- Skarnitzl, R. (2011). *Znělostní kontrast nejen v češtině*. Praha: Nakladatelství Epocha.
- Skarnitzl, R. & Machač, P. (2012). Míra rušivosti parazitních zvuků v řeči mediálních mluvčích. *Naše řeč*, 95, 3–14.
- Sluijter, A. & van Heuven, V. (1996a). Acoustic correlates of linguistic stress and accent in Dutch and American English. In: *Proceedings of the International Congress of Spoken Language Processing*, Philadelphia, 630–633.
- Sluijter, A. & van Heuven, V. (1996b). Spectral balance as an acoustic correlate of linguistic stress. *Journal of the Acoustical Society of America*, 100, 2471–2485.
- Solé, M. J. & Ohala, J. J. (2010). What is and what is not under the control of the speaker. Intrinsic vowel duration. In: Fougeron, C., Kühnert, B., D'Imperio, M. & Vallée, N. (Eds.), *Papers in Laboratory Phonology 10*, 607–655. Berlin: de Gruyter.
- Studenovský, D. (2012). *Akustické vlastnosti českých diftongů*. Disertační práce. Praha: Univerzita Karlova.
- Šimko, J., Aalto, D., Lippus, P., Włodarczak, M. & Vainio, M. (2015). Pitch, perceived duration and auditory biases: comparison among languages. In: *Proceedings of the 18<sup>th</sup> ICPHS*. Glasgow: University of Glasgow.
- Šturm, P. & Volín, J. (2016). P-centres in natural disyllabic Czech words in a large-scale speech-metronome synchronization experiment. *Journal of Phonetics*, 55, 38–52.
- Trávníček, F. (1932). *Úvod do české fonetiky*. Praha: Česká grafická Unie.
- Van Santen, J. P. H. (1992). Contextual effects on vowel duration. *Speech Communication*, 11, 513–546.
- Veroňková, J. (2012). Tempo řeči z různých stran. In: Hajíčková, Z. & Vacula, R. (Eds.), *Sborník Asociace učitelů češtiny jako cizího jazyka (AUČCJ) 2012*, 203–222. Praha: Filip Tomáš – Akropolis.
- Veroňková, J. & Janoušková, J. (2010). Jak hodnotí posluchači zvukovou stránku projevů u moderátorů zpravodajství ČT1? In: Čmejrková, S., Havlová, E. & Hoffmannová, J. (Eds.), *Užívání a prožívání jazyka*, 115–119. Praha: Karolinum.

- Volín, J. (2009). Extrakce základní hlasové frekvence a intonační gravitace v češtině. *Naše řeč*, 92(5), 227–239.
- Volín, J. (2009). Metric warping in Czech newsreading. In: Vích, R. (Ed.), *Speech Processing – 19th Czech-German Workshop*, 52–55. Praha.
- Volín, J. & Churaňová, E. (2010). Probabilities of consonantal sequences in continuous Czech texts. *AUC Philologica 1, Phonetica Pragensia XII*, 49–62.
- Volín, J. & Weingartová, L. (2014). Acoustic correlates of word stress as a cue to accent strength. *Research in Language*, 12(2), 175–183.
- Volín, J., Churaňová, E. & Šturm, P. (2014). P-centre position in natural two-syllable Czech words. In: *Proceedings of Speech Prosody 2014*, 920–924. Dublin: Trinity College Dublin.
- Weingartová, L. (2015). *Identifikace mluvčího v temporální doméně řeči*. Disertační práce. Praha: Univerzita Karlova.
- Zima, P. (1959). K otázce klasifikace mluvního tempa. *Slovo a slovesnost*, 20, 96–117.

## Přílohy

### Příloha 1 – položky testu

Rozpis všech položek v percepčním experimentu. Ve sloupci *položky* jsou čísla analyzovaných položek. Položky s označením Px obsahují rozdílné takty nebo takt a nízkofrekvenční obraz jiného taktu, položky s označením Nx obsahují takt a nízkofrekvenční obraz téhož taktu. Označení položky je psáno kurzívou, pokud v položce byl manipulován průběh intonace. Ve sloupci *takt 1* je uveden slovní obsah prvního taktu položky; sloupec *takt 2* zobrazuje obsah druhého taktu v položce. Text v tomto sloupci psaný kurzívou značí, že druhý takt v položce zazněl jen jako nízkofrekvenční stín. V dalších sloupcích jsou zapsány fonotaktické vzorce daných taktů, kde S = sonora, Z = znělý obstruent, N = neznělý obstruent, V = vokál. V posledním sloupci je uveden celkový počet hodnocení dané položky (odpovídá počtu výskytů v jednotlivých variantách testu: 40 = ve všech variantách 1×, 80 = ve všech variantách 2×, 13 a 14 = v jedné variantě 1×).

položky	takt 1	takt 2	vzorec 1	vzorec 2	počet hodnocení
P1	fáze	<i>jsou to</i>	NVZV	NVNV	40
P2	státy	<i>svého</i>	NNVNV	NZVZV	40
P3	kdy se	<i>škody</i>	ZZVNV	NNVZV	40
P4	se zde	<i>Šopko</i>	NVZZV	NVNNV	40
P5	mezi	<i>jako</i>	SVZV	SVNV	80
P6	jsou to	<i>žáci</i>	NVNV	ZVNV	40
P7	škody	<i>v Gaze</i>	NNVZV	ZZVZV	40
P8	fáze	<i>žáci</i>	NVZV	ZVNV	40
P9	kdy se	<i>v Gaze</i>	ZZVNV	ZZVZV	40
P10	bude	<i>v Gaze</i>	ZVZV	ZZVZV	40
P11	jsou to	<i>státy</i>	NVNV	NNVNV	40
P12	byli	<i>sboru</i>	ZVSV	ZZVSV	80
P13	žáci	<i>výšce</i>	ZVNV	ZVNNV	40
P14	mezi	<i>na dva</i>	SVZV	SVZZV	40
P15	nimi	<i>země</i>	SVSV	ZVSSV	40
P16	bude	<i>země</i>	ZVZV	ZVSSV	40
P17	dále	<i>mírně</i>	ZVSV	SVSSV	80
P18	dále	<i>jméno</i>	ZVSV	SSVSV	40
P19	fáze	<i>stále</i>	NVZV	NNVSV	40
P20	jméno	<i>mírně</i>	SSVSV	SVSSV	40
P21	vjely	<i>země</i>	ZSVSV	ZVSSV	40
P22	sboru	<i>době</i>	ZZVSV	ZVZSV	40

P23	sešlo	škody	NVNSV	NNVZV	40
P24	vjely	volby	ZSVSV	ZVSZV	40
P25	vjely	jedna	ZSVSV	SVZSV	40
P26	vjely	době	ZSVSV	ZVZSV	40
P27	sboru	jedna	ZZVSV	SVZSV	40
P28	sboru	volby	ZZVSV	ZVSZV	40
P29	sboru	země	ZZVSV	ZVSSV	40
P30	bude	nimi	ZVZV	SVSV	40
P31	jedna	době	SVZSV	ZVZSV	40
P32	vjely	sboru	ZSVSV	ZZVSV	40
P33	vládě	jméno	ZSVZV	SSVSV	40
P34	v noci	Lhase	ZSVNV	SZVNV	40
P35	výšce	místy	ZVNNV	SVNNV	40
P36	se zde	silně	NVZZV	NVSSV	40
P37	nimi	byli	SVSV	ZVSV	40
P38	byli	mezi	ZVSV	SVZV	40
P39	bude	byli	ZVZV	ZVSV	40
P40	bude	mezi	ZVZV	SVZV	40
P41	nimi	mezi	SVSV	SVZV	40
P42	jsou to	fáze	NVNV	NVZV	40
P43	svého	státy	NZVZV	NNVNV	40
P44	škody	kdy se	NNVZV	ZZVNV	40
P45	Šopko	se zde	NVNNV	NVZZV	80
P46	jako	mezi	SVNV	SVZV	40
P47	žáci	jsou to	ZVNV	NVNV	40
P48	v Gaze	škody	ZZVZV	NNVZV	40
P49	žáci	fáze	ZVNV	NVZV	40
P50	v Gaze	kdy se	ZZVZV	ZZVNV	40
P51	v Gaze	bude	ZZVZV	ZVZV	40
P52	státy	jsou to	NNVNV	NVNV	40
P53	sboru	byli	ZZVSV	ZVSV	40
P54	výšce	žáci	ZVNNV	ZVNV	40
P55	na dva	mezi	SVZZV	SVZV	40
P56	země	nimi	ZVSSV	SVSV	40
P57	země	bude	ZVSSV	ZVZV	40
P58	mírně	dále	SVSSV	ZVSV	40
P59	jméno	dále	SSVSV	ZVSV	40
P60	stále	fáze	NNVSV	NVZV	40
P61	mírně	jméno	SVSSV	SSVSV	40
P62	země	vjely	ZVSSV	ZSVSV	40
P63	době	sboru	ZVZSV	ZZVSV	40
P64	škody	sešlo	NNVZV	NVNSV	80
P65	volby	vjely	ZVSZV	ZSVSV	40
P66	jedna	vjely	SVZSV	ZSVSV	40

P67	době	<i>vjely</i>	ZVZSV	ZSVSV	40
P68	jedna	<i>sboru</i>	SVZSV	ZZVSV	40
P69	volby	<i>sboru</i>	ZVSZV	ZZVSV	40
P70	země	<i>sboru</i>	ZVSSV	ZZVSV	40
P71	nimi	<i>bude</i>	SVSV	ZVZV	40
P72	době	<i>jedna</i>	ZVZSV	SVZSV	40
P73	sboru	<i>vjely</i>	ZZVSV	ZSVSV	40
P74	jméno	<i>vládě</i>	SSVSV	ZSVZV	80
P75	Lhase	<i>v noci</i>	SZVNV	ZSVNV	40
P76	místy	<i>výšce</i>	SVNNV	ZVNNV	40
P77	silně	<i>se zde</i>	NVSSV	NVZZV	40
P78	byli	<i>nimi</i>	ZVSV	SVSV	40
P79	mezi	<i>byli</i>	SVZV	ZVSV	40
P80	byli	<i>bude</i>	ZVSV	ZVZV	40
P81	mezi	<i>bude</i>	SVZV	ZVZV	40
P82	mezi	<i>nimi</i>	SVZV	SVSV	40
P83	fáze	<i>jsou to</i>	NVZV	NVNV	40
P84	státy	<i>svého</i>	NNVNV	NZVZV	40
P85	kdy se	<i>škody</i>	ZZVNV	NNVZV	40
P86	se zde	<i>Šopko</i>	NVZZV	NVNNV	40
P87	mezi	<i>jako</i>	SVZV	SVNV	40
P88	jsou to	<i>žáci</i>	NVNV	ZVNV	40
P89	škody	<i>v Gaze</i>	NNVZV	ZZVZV	40
P90	fáze	<i>žáci</i>	NVZV	ZVNV	40
P91	kdy se	<i>v Gaze</i>	ZZVNV	ZZVZV	40
P92	bude	<i>v Gaze</i>	ZVZV	ZZVZV	40
P93	jsou to	<i>státy</i>	NVNV	NNVNV	40
P94	byli	<i>sboru</i>	ZVSV	ZZVSV	40
P95	žáci	<i>výšce</i>	ZVNV	ZVNNV	80
P96	mezi	<i>na dva</i>	SVZV	SVZZV	40
P97	nimi	<i>země</i>	SVSV	ZVSSV	40
P98	bude	<i>země</i>	ZVZV	ZVSSV	40
P99	dále	<i>mírně</i>	ZVSV	SVSSV	40
P100	dále	<i>jméno</i>	ZVSV	SSVSV	40
P101	fáze	<i>stále</i>	NVZV	NNVSV	40
P102	jméno	<i>mírně</i>	SSVSV	SVSSV	40
P103	vjely	<i>země</i>	ZSVSV	ZVSSV	40
P104	sboru	<i>době</i>	ZZVSV	ZVZSV	40
P105	sešlo	<i>škody</i>	NVNSV	NNVZV	40
P106	vjely	<i>volby</i>	ZSVSV	ZVSZV	40
P107	vjely	<i>jedna</i>	ZSVSV	SVZSV	40
P108	vjely	<i>době</i>	ZSVSV	ZVZSV	40
P109	sboru	<i>jedna</i>	ZZVSV	SVZSV	40
P110	sboru	<i>volby</i>	ZZVSV	ZVSZV	40



P111	sboru	země	ZZVSV	ZVSSV	40
P112	bude	nimi	ZVZV	SVSV	40
P113	jedna	době	SVZSV	ZVZSV	40
P114	vjely	sboru	ZSVSV	ZZVSV	40
P115	vládě	jméno	ZSVZV	SSVSV	40
P116	v noci	Lhase	ZSVNV	SZVNV	40
P117	výšce	místy	ZVNNV	SVNNV	40
P118	se zde	silně	NVZZV	NVSSV	40
P119	nimi	byli	SVSV	ZVSV	40
P120	byli	mezi	ZVSV	SVZV	40
P121	bude	byli	ZVZV	ZVSV	40
P122	bude	mezi	ZVZV	SVZV	40
P123	nimi	mezi	SVSV	SVZV	80
P124	jsou to	fáze	NVNV	NVZV	40
P125	svého	státy	NZVZV	NNVNV	40
P126	škody	kdy se	NNVZV	ZZVNV	40
P127	Šopko	se zde	NVNNV	NVZZV	40
P128	jako	mezi	SVNV	SVZV	40
P129	žáci	jsou to	ZVNV	NVNV	40
P130	v Gaze	škody	ZZVZV	NNVZV	40
P131	žáci	fáze	ZVNV	NVZV	40
P132	v Gaze	kdy se	ZZVZV	ZZVNV	40
P133	v Gaze	bude	ZZVZV	ZVZV	40
P134	státy	jsou to	NNVNV	NVNV	40
P135	sboru	byli	ZZVSV	ZVSV	40
P136	výšce	žáci	ZVNNV	ZVNV	40
P137	na dva	mezi	SVZZV	SVZV	40
P138	země	nimi	ZVSSV	SVSV	40
P139	země	bude	ZVSSV	ZVZV	40
P140	mírně	dále	SVSSV	ZVSV	40
P141	jméno	dále	SSVSV	ZVSV	40
P142	stále	fáze	NNVSV	NVZV	40
P143	mírně	jméno	SVSSV	SSVSV	40
P144	země	vjely	ZVSSV	ZSVSV	40
P145	době	sboru	ZVZSV	ZZVSV	40
P146	škody	sešlo	NNVZV	NVNSV	40
P147	volby	vjely	ZVSZV	ZSVSV	80
P148	jedna	vjely	SVZSV	ZSVSV	40
P149	době	vjely	ZVZSV	ZSVSV	40
P150	jedna	sboru	SVZSV	ZZVSV	40
P151	volby	sboru	ZVSZV	ZZVSV	40
P152	země	sboru	ZVSSV	ZZVSV	40
P153	nimi	bude	SVSV	ZVZV	40
P154	době	jedna	ZVZSV	SVZSV	40

P155	sboru	vjely	ZZVSV	ZSVSV	40
P156	jméno	vládě	SSVSV	ZSVZV	40
P157	Lhase	v noci	SZVNV	ZSVNV	40
P158	místy	výšce	SVNNV	ZVNNV	80
P159	silně	se zde	NVSSV	NVZZV	40
P160	byli	nimi	ZVSV	SVSV	40
P161	mezi	byli	SVZV	ZVSV	40
P162	byli	bude	ZVSV	ZVZV	40
P163	mezi	bude	SVZV	ZVZV	40
P164	mezi	nimi	SVZV	SVSV	40
P165	mezi	jako	SVZV	SVNV	80
P166	mezi	na dva	SVZV	SVZZV	40
P167	sboru	době	ZZVSV	ZVZSV	40
P168	vládě	jméno	ZSVZV	SSVSV	40
P169	výšce	místy	ZVNNV	SVNNV	40
P170	jako	mezi	SVNV	SVZV	40
P171	na dva	mezi	SVZZV	SVZV	40
P172	době	sboru	ZVZSV	ZZVSV	40
P173	jméno	vládě	SSVSV	ZSVZV	80
P174	místy	výšce	SVNNV	ZVNNV	40
P175	mezi	jako	SVZV	SVNV	40
P176	mezi	na dva	SVZV	SVZZV	40
P177	sboru	době	ZZVSV	ZVZSV	40
P178	vládě	jméno	ZSVZV	SSVSV	40
P179	výšce	místy	ZVNNV	SVNNV	40
P180	jako	mezi	SVNV	SVZV	40
P181	na dva	mezi	SVZZV	SVZV	80
P182	době	sboru	ZVZSV	ZZVSV	40
P183	jméno	vládě	SSVSV	ZSVZV	40
P184	místy	výšce	SVNNV	ZVNNV	40
N1	fáze	fáze	NVZV	NVZV	13
N2	státy	státy	NNVNV	NNVNV	13
N3	kdy se	kdy se	ZZVNV	ZZVNV	13
N4	se zde	se zde	NVZZV	NVZZV	13
N5	mezi	mezi	SVZV	SVZV	13
N6	škody	škody	NNVZV	NNVZV	13
N7	bude	bude	ZVZV	ZVZV	13
N8	jsou to	jsou to	NVNV	NVNV	13
N9	byli	byli	ZVSV	ZVSV	13
N10	žáci	žáci	ZVNV	ZVNV	13
N11	nimi	nimi	SVSV	SVSV	13
N12	dále	dále	ZVSV	ZVSV	13
N13	jméno	jméno	SSVSV	SSVSV	13
N14	vjely	vjely	ZSVSV	ZSVSV	14

N15	sboru	<i>sboru</i>	ZZVSV	ZZVSV	14
N16	sešlo	<i>sešlo</i>	NVNSV	NVNSV	14
N17	jedna	<i>jedna</i>	SVZSV	SVZSV	14
N18	vládě	<i>vládě</i>	ZSVZV	ZSVZV	14
N19	v noci	<i>v noci</i>	ZSVNV	ZSVNV	14
N20	výšce	<i>výšce</i>	ZVNNV	ZVNNV	14
N21	svého	<i>svého</i>	NZVZV	NZVZV	14
N22	Šopko	<i>Šopko</i>	NVNNV	NVNNV	14
N23	jako	<i>jako</i>	SVNV	SVNV	14
N24	v Gaze	<i>v Gaze</i>	ZZVZV	ZZVZV	14
N25	na dva	<i>na dva</i>	SVZZV	SVZZV	14
N26	země	<i>země</i>	ZVSSV	ZVSSV	14
N27	stále	<i>stále</i>	NNVSV	NNVSV	13
N28	volby	<i>volby</i>	ZVSZV	ZVSZV	13
N29	době	<i>době</i>	ZVZSV	ZVZSV	13
N30	Lhase	<i>Lhase</i>	SZVNV	SZVNV	13
N31	místy	<i>místy</i>	SVNNV	SVNNV	13
N32	silně	<i>silně</i>	NVSSV	NVSSV	13
N33	mezi	<i>mezi</i>	SVZV	SVZV	13
N34	sboru	<i>sboru</i>	ZZVSV	ZZVSV	13
N35	vládě	<i>vládě</i>	ZSVZV	ZSVZV	13
N36	výšce	<i>výšce</i>	ZVNNV	ZVNNV	13
N37	jako	<i>jako</i>	SVNV	SVNV	13
N38	na dva	<i>na dva</i>	SVZZV	SVZZV	13
N39	době	<i>době</i>	ZVZSV	ZVZSV	13

## Příloha 2 – přepisy jednotlivých mluvních taktů v testu

text	IPA
fáze	fa:zɛ
státy	sta:tɪ
kdy se	gdɪ sɛ
se zde	sɛ zɛɛ
mezi	mɛzɪ
jsou to	sou to
škody	ʃkɔdɪ
bude	budɛ
byli	bɪlɪ
žáci	ʒa:tɪ
nimi	nɪmɪ
dále	da:lɛ
jméno	jmɛ:no
vjely	vjelɪ
sboru	zboru
sešlo	sɛflo

jedna	jedna
vládě	vla:ʒɛ
v noci	v notsɪ
výšce	vi:ʃtsɛ
svého	svɛ:fo
Šopko	ʃopko
jako	jako
v Gaze	v gaze
na dva	na dva
země	zɛmɲɛ
mírně	mi:rɲɛ
stále	sta:lɛ
době	dobɛ
volby	volbɪ
Lhase	lʰase
místy	mi:stɪ
silně	silɲɛ